

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 16:38:18 ; Search time 1625 Seconds  
(without alignments)  
3222.420 Million cell updates/sec

Title: US-09-889-611-1  
Perfect score: 128  
Sequence: 1 aatgaactacataacaacca.....gaagccatcccaagaagccag 128

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.nam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	123.2	96.2	4229	9	AF234618	AF234618 Homo sapi
2	123.2	96.2	157284	9	AC072051	AC072051 Homo sapi
3	123.2	96.2	189092	2	AC027524	AC027524 Homo sapi
4	85.4	66.7	6182	6	AX346917	AX346917 Sequence
5	78.4	61.3	6182	6	AX346916	AX346916 Sequence
6	60.4	47.2	199669	10	AC103453	AC103453 Rattus no
7	60.4	47.2	219071	2	AC133259	AC133259 Rattus no
8	58.8	45.9	128075	2	AC129295	AC129295 Mus muscu
9	35.6	27.8	237282	2	AC112948	AC112948 Mus muscu
10	34.8	27.2	165860	10	AL831716	AL831716 Mouse DNA
11	34.8	27.2	251504	2	AC131633	AC131633 Rattus no
12	34.4	26.9	110000	2	HSY31394_1	Continuation (2 of
13	34.4	26.9	158806	2	AL591132	Continuation (2 of
14	34.4	26.9	168834	2	AL355583	AL591132 Homo sapi
15	34.4	26.9	194181	9	AC124148	AL355583 Homo sapi
16	34.4	26.9	197864	9	HSY214H10	AC124148 Pan trogl
17	34.4	26.9	222728	2	AC141289	AL022344 Human DNA
18	34.4	26.9	238379	2	AL590986	AC141289 Homo sapi
19	33.8	26.4	212670	2	EX005210	AL590986 Homo sapi
20	33.8	26.4	229529	2	AC095130	EX005210 Danilo ter
21	33.8	26.4	251779	2	AC132691	AC095130 Rattus no
22	33.2	25.9	101765	9	AL136306	AC132691 Rattus no
23	33.2	25.9	143526	2	AC009650	AL136306 Human DNA
24	33.2	25.9	168141	2	AC137465	AC009650 Homo sapi
25	33.2	25.9	179356	2	AC121403	AC137465 Rattus no
26	33.2	25.9	181786	2	AC119376	AC121403 Rattus no
27	33.2	25.9	219949	2	AC119687	AC119376 Rattus no
28	33	25.8	220643	2	AC113804	AC119687 Rattus no
29	33	25.8	293021	2	AC098063	AC113804 Rattus no
30	32.8	25.6	59105	9	AC004013	AC098063 Rattus no
31	32.8	25.6	110000	2	AC120731_2	AC004013 Homo sapi
32	32.8	25.6	116859	9	AL358394	Continuation (3 of
33	32.8	25.6	128615	9	AL591051	Continuation (3 of
34	32.8	25.6	143893	2	AC010766	AL358394 Human DNA
35	32.8	25.6	155629	9	AC125391	AL591051 Human DNA
36	32.8	25.6	162209	9	CNS01RIH	AC010766 Homo sapi
37	32.8	25.6	170008	9	HS010770	AC125391 Pan trogl
38	32.8	25.6	170278	2	AC020586	AL163195 Human chr
39	32.8	25.6	173540	2	AC128242	AC020586 Homo sapi
40	32.8	25.6	181603	2	AC093189	AC128242 Rattus no
41	32.8	25.6	245426	2	AC103108	AC093189 Pan trogl
42	32.8	25.6	251872	2	AC094163	AC103108 Rattus no
43	32.8	25.6	254641	2	AC103412	AC094163 Rattus no
44	32.8	25.6	294459	2	AC128978	AC103412 Rattus no
45	32.6	25.5	345109	2	AC098623	AC128978 Rattus no

ALIGNMENTS

RESULT 1  
AP234618  
LOCUS AF234618 4229 bp DNA linear PRI 29-MAR-2002  
DEFINITION Homo sapiens megsin gene, promoter region and partial sequence.  
ACCESSION AF234618  
VERSION AF234618.2 GI:19808130  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 4229)  
REFERENCE  
AUTHORS Miyata.T., Nangaku.M., Inagi.R. and Kurokawa.K.  
TITLE Transcriptional regulation of a mesangium-predominant gene, megsin  
JOURNAL Unpublished

Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Raymond, J., Peterson, K., Pierre, N., Pisaní, C., Pollara, V., Chango, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrelli, A., Travers, M., Triggilo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (07-JUN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (Bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lanazares, R., Landers, T., Lehoczeky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McGowan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneilly, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (25-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (Bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lanazares, R., Landers, T., Lehoczeky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McGowan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneilly, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (03-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Nov 18, 2001 this sequence version replaced gi:14277307.

All repeats were identified using RepeatMasker:  
Smt, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information

Center project name: L966  
Center clone name: 79 D 21

## FEATURES

## source

## Location/Qualifiers

source	1. .157284	
repeat_region	/mol_type="Homo sapiens"	
repeat_region	/db_xref="taxon:9606"	
repeat_region	/clone="RP11-79D21"	
repeat_region	/clone_lib="RPC1-11 Human Male BAC"	
repeat_region	599. .635	
repeat_region	/rpt_family="(T)n"	
repeat_region	1040. .1077	
repeat_region	/rpt_family="AT_rich"	
repeat_region	1411. .1747	
repeat_region	/rpt_family="L1p"	
repeat_region	2632. .2992	
repeat_region	/rpt_family="MLT1A1"	
repeat_region	3115. .3138	
repeat_region	/rpt_family="AT_rich"	
repeat_region	complement(3487. .3828)	
repeat_region	/rpt_family="L1MC4a"	
repeat_region	3834. .4338	
repeat_region	/rpt_family="MLT1C"	
repeat_region	complement(4339. .5397)	
repeat_region	/rpt_family="L1MC4a"	
repeat_region	complement(5742. .5899)	
repeat_region	/rpt_family="MER5B"	
repeat_region	6499. .6816	
repeat_region	/rpt_family="HALL1"	
repeat_region	complement(7416. .7621)	
repeat_region	/rpt_family="MER3"	
repeat_region	complement(8163. .9090)	
repeat_region	/rpt_family="L2"	
repeat_region	9276. .9419	
repeat_region	/rpt_family="MIR"	
repeat_region	9598. .9658	
repeat_region	/rpt_family="MER381"	
repeat_region	9659. .10033	
repeat_region	/rpt_family="MLT1J"	
repeat_region	13986. .14009	
repeat_region	/rpt_family="AT_rich"	
repeat_region	14011. .14115	
repeat_region	/rpt_family="(TA)n"	
repeat_region	complement(14942. .15044)	
repeat_region	/rpt_family="MIR3"	
repeat_region	15156. .15298	
repeat_region	/rpt_family="L2"	
repeat_region	15592. .15773	
repeat_region	/rpt_family="L1W4C"	
repeat_region	15874. .18310	
repeat_region	/rpt_family="L1PAL6"	
unsure	16373. .16374	
unsure	/note="1328 bp of bacterial transposon excised between these 2 bp"	
repeat_region	18311. .18331	
repeat_region	/rpt_family="(TAAAA)n"	
repeat_region	18332. .19566	
repeat_region	/rpt_family="L1PAL6"	
repeat_region	19567. .19587	
repeat_region	/rpt_family="(CA)n"	
repeat_region	19588. .20042	
repeat_region	/rpt_family="L1PAL6"	
repeat_region	20043. .20064	
repeat_region	/rpt_family="AT_rich"	
repeat_region	complement(21259. .21386)	
repeat_region	/rpt_family="MIR3"	
repeat_region	22395. .22420	
repeat_region	/rpt_family="AT_rich"	
repeat_region	complement(22788. .22948)	
repeat_region	/rpt_family="FRAM"	
repeat_region	complement(23843. .24253)	
repeat_region	/rpt_family="MSTA"	

repeat_region	complement (246758, .25041)
repeat_region	/rpt family="MUT1A2"
repeat_region	25054, .25094
repeat_region	/rpt family="HAL1"
repeat_region	25214, .25358
repeat_region	/rpt family="LIM4C"
repeat_region	complement (25602, .25854)
repeat_region	/rpt family="MIR"
repeat_region	26746, .26796
repeat_region	/rpt family="AT_rich"
repeat_region	27451, .27494
repeat_region	/rpt family=" (CAAAA)"n"
repeat_region	complement (28193, .28219)
repeat_region	/rpt family="MIR"
repeat_region	complement (28249, .28325)
repeat_region	/rpt family="MIR"
repeat_region	complement (29372, .29459)
repeat_region	/rpt family="MIR"
repeat_region	29539, .30003
repeat_region	/rpt family="MER66B"
repeat_region	30596, .30624
repeat_region	/rpt family="AT_rich"
repeat_region	31418, 31829
repeat_region	/rpt family="I2"
repeat_region	31846, .31915
repeat_region	/rpt family=" (TGAA)"n"
repeat_region	complement (32019, .32103)

Query Match 96.2%; Score 123.2; DB 9; Length 157284;  
Best Local Similarity 97.7%; Pred. No. 3.8e-25;  
Matches 125; Conservative 0; Mismatches 3; Indels 0;

QY	1	AATCAAGCTCATTAACACCACTTAGTCAAGTACTCTTTGAACCTCGTTTCAAAACCTA	60
Db	155066	AATGACTCATTAACACCACTTAGTCAAGTACTCTTTGAACCTCGTTTCAAAACCTA	155007
QY	61	AATGCTTATAAGARRCTTGAGACAGCTGTGTGCTCTGAGTCATAGGGAAGCCATCCCA	120
Db	155006	AATGCTTATAAGATCTTGAGACAGCTGTGTGCTCTGAGTCATAGGGAAGCCATCCCA	154947
QY	121	GAAGCCAG	128
Db	154946	GAAGCCAG	154939

RESIT, T 3

AC027524 189092 bp DNA linear HTG 27-MAR-2003  
 LOCUS Homo sapiens chromosome 18 clone RP11-317G1 map 18, 3 unordered  
 DEFINITION pieces.  
 ACCESSION AC027524 GI:21307437  
 VERSION HTG, HTGS\_PHASE1, HTGS\_FULLTOP; HTGS\_CANCELLED.  
 KEYWORDS Homo sapiens  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 18, clone RP11-317G1  
 Unpublished  
 2 (bases 1 to 189092)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
 Boguslavsky,L., Bouckaiter,B., Brown,A., Burkett,G.,  
 Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
 Grand-Pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
 Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Manuif,N.,

McCarthy M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (30-MAR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 189092)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukagater, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, F., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,  
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,  
Lamarez, R., Landers, R., Lehoczy, J., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N.,  
Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,  
Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,  
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,  
Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,  
Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zembek, B., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 1, 2002 this sequence version replaced gi:11990731.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information  
Center project name: 18516  
Center clone name: 317\_G\_1

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 85659: contig of 85659 bp in length  
\* 85660 85759: gap of 100 bp  
\* 85760 163014: contig of 77255 bp in length  
\* 163015 163114: gap of 100 bp  
\* 163115 189092: contig of 25978 bp in length.

Location/Qualifiers  
1. .189092  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18"  
/clone="RP11-317G1"

FEATURES  
source

BASE COUNT 57107 a 35683 c 36719 g 59369 t 214 others  
ORIGIN  
Query Match 96.2%; Score 123.2; DB 2; Length 189092;  
Best Local Similarity 97.7%; Pred. No. 3.7e-25;  
Matches 125; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 AATGAACACTACATAACCAACACCTTAGTCAGATACACTTTGAAACCTGGTTCAAAACCTA 60  
DB 66564 AATGAACACTACATAACCAACACCTTAGTCAGATACACTTTGAAACCTGGTTCAAAACCTA 66623  
QY 61 AATGCTTATAAGARCTTTGAGAGACAGTGTGTCTCTAGTCATAGGAGGAGCCATCCCA 120  
DB 66624 AATGCTTATAAGTCTTTGAGAGACAGTGTGTCTCTAGTCATAGGAGGAGCCATCCCA 66683  
QY 121 GAAGCCAG 128  
DB 66684 GAAGCCAG 66691  
RESULT 4  
AX346917/c AX346917 6182 bp DNA linear PAT 01-FEB-2002  
LOCUS  
DEFINITION Sequence 1988 from Patent WO0200928.  
ACCESSION AX346917  
VERSION AX346917.1 GI:18494803  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with the immune system  
JOURNAL Patent: WO 0200928-A 1988 03-JAN-2002;  
Epidemiology AG (DE)  
FEATURES  
Location/Qualifiers  
1. .6182  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 1934 a 37 c 1059 g 3152 t  
ORIGIN  
Query Match 66.7%; Score 85.4; DB 6; Length 6182;  
Best Local Similarity 79.5%; Pred. No. 3.7e-14;  
Matches 101; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 AATGAACACTACATAACCAACACCTTAGTCAGATACACTTTGAAACCTGGTTCAAAACCTA 60  
DB 1310 AATGAACACTACATAACCAACACCTTAGTCAGATACACTTTGAAACCTGGTTCAAAACCTA 1251  
QY 61 AATGCTTATAAGARCTTTGAGAGACAGTGTGTCTCTAGTCATAGGAGGAGCCATCCCA 120  
DB 1250 AATGCTTATAAGTCTTTGAGAGACAGTGTGTCTCTAGTCATAGGAGGAGCCATCCCA 1191  
QY 121 GAAGCCCA 127  
DB 1190 AAACCA 1184  
RESULT 5  
AX346916  
LOCUS  
DEFINITION Sequence 1987 from Patent WO0200928.  
ACCESSION AX346916  
VERSION AX346916.1 GI:18494802  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1

Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwundu, G., Olanrunpaagoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C.D., Smajd, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J., Steagle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Wadron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlecsyk, R., Wooden, H., Worley, K., Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 199669)  
Worley, K.C.

Direct Submission  
Submitted (25-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 199669)

Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 199669)  
Worley, K.C.

Direct Submission  
Submitted (07-JUN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jun 7, 2003 this sequence version replaced gi:30521373.  
Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

Location/Qualifiers  
1..199669  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/chromosome="13"  
/clone="CH230-127N19"  
\_region 5394..5759  
/rpt\_family="Lx2\_3"  
\_region 5760..5861  
/rpt\_family="(CA)n"  
\_region 5862..6447  
/rpt\_family="Lx2\_3"  
\_region 6522..6586  
/rpt\_family="(TG)n"  
\_region 6611..6938  
/rpt\_family="L1M1"  
\_region 6939..7037  
/rpt\_family="(TA)n"  
\_region complement(7044..7089)  
/rpt\_family="FMR1B"  
\_region 7090..7169  
/rpt\_family="(TCTA)n"  
\_region complement(7190..7799)  
/rpt\_family="FMR1B"  
\_region complement(9049..10178)  
/rpt\_family="L1"  
\_region complement(10334..10982)









Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (12-JUN-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jun 12, 2003 this sequence version replaced gi:30985018.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L20361  
 Center clone name: 299\_1\_18  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 28064: contig of 28064 bp in length  
 \* 28065 28164: gap of 100 bp  
 \* 28165 32220: contig of 7056 bp in length  
 \* 32221 35321: gap of 100 bp  
 \* 35321 237282: contig of 201962 bp in length.  
 \* Location/Qualifiers  
 1. 237282  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP23-299118"  
 /clone\_lib="RPCI-23 Female Mouse BAC"  
 BASE COUNT 73426 a 44157 c 43655 g 75844 t 200 others  
 ORIGIN  
 Query Match 27.8%; Score 35.6; DB 2; Length 237282;  
 Best Local Similarity 57.8%; Pred. No. 5;  
 Matches 59; Conservative 2; Mismatches 41; Indels 0; Gaps 0;  
 QY 1 ATGACATACATACACACCTTGTGACATCTTGTGACACCTGTTCAAACTTA 60  
 Db 195228 ATTAACCTTCTTAAACAGCTGATGAGTTAAATCTGTGCGACCTTTGACACAGCCCA 195287  
 QY 61 AATGCTTATAAGARRCTTGACACAGTGTCTGTGCTCTGAGT 102  
 Db 195288 ATAGCTTTTACAGTTTGACAGATTAAGCTGAGCTGTGTAAT 195329  
 -----  
 RESULT 10  
 AL831716 165860 bp DNA linear ROD 11-APR-2003  
 LOCUS  
 DEFINITION Mouse DNA sequence from clone RP23-280E8 on chromosome X, complete  
 sequence.  
 ACCESSION AL831716  
 VERSION AL831716.5 GI:29823187  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 165860)  
 Clark,S  
 Direct Submission  
 Submitted (11-APR-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

## COMMENT

On Apr 11, 2003 this sequence version replaced gi:22859065.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
 was finished as follows unless otherwise noted: all regions were  
 either double-stranded or sequenced with an alternate chemistry or  
 covered by high quality data (i.e., phred quality >= 30); an  
 attempt was made to resolve all sequencing problems, such as  
 compressions and repeats; all regions were covered by at least one  
 plasmid subclone or more than one M13 subclone; and the assembly  
 was confirmed by restriction digest, except on the rare occasion of  
 the clone being a YAC.  
 RP23-280E8 is from the RPCI-23 Mouse BAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see http://www.chori.org/bacpac/home.htm  
 VECTOR: pBAC3.6  
 Sequence from the Mouse Genome Sequencing Consortium whole genome  
 shotgun may have been used to confirm this sequence. Sequence data  
 from the whole genome shotgun alone has only been used where it has  
 a phred quality of at least 30.

FEATURES  
source

1. 165860  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP23-280E8"  
 /clone\_lib="RPCI-23"  
 BASE COUNT 49259 a 31599 c 32303 g 52699 t  
 ORIGIN

Query Match 27.2%; Score 34.8; DB 10; Length 165860;  
 Best Local Similarity 78.0%; Pred. No. 8.8;  
 Matches 39; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 52 CAAACCTTAATGCTTATAAGARRCTTGAGACAGTGTGTGCTCTGAG 101  
 Db 24443 CAAACCTTAACGATATAAAGCTTGAGATCAGAGTTGAACTCTGAG 24492

RESULT 11  
AC131633/c

LOCUS AC131633 251504 bp DNA linear HTG 10-MAY-2003  
 DEFINITION Rattus norvegicus clone CH230-11K20, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 2 unordered pieces.  
 ACCESSION AC131633  
 VERSION AC131633.4 GI:30522901  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 251504)

REFERENCE  
AUTHORS

Muzny,D,Marie, Metzker,M,Lee, Abramson,S., Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,  
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, B., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Haravey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Huly, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, H., Loulsegged, H., Lozano, R.J., Lu, X., Ma, J., Maneshwari, M., Maindardne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokeneme, O., Okwuonu, G., Olarunpugoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, H., Pfankuch, C., Plummer, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2 (bases 1 to 251504)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (25-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 251504)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:24941108.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GEBZ  
Center clone name: CH230-11K20  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 231072 bases at least Q40  
Consensus quality: 234181 bases at least Q30  
Consensus quality: 236130 bases at least Q20  
Estimated insert size: 244864; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 249063: contig of 249063 bp in length  
\* 249064 249163: gap of unknown length  
\* 249164 251504: contig of 2341 bp in length.

FEATURES

source

Location/Qualifiers

1..251504

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-11K20"

1..2016

/note="wgs\_end\_extension"

clone\_end:T7"

16230..16885

/note="clone\_boundary"

clone\_end:T7"

site:EcoRI

end\_sequence:BH342746"

complement(243291..244225)

/note="clone\_boundary"

clone\_end:Sp6"

site:EcoRI

end\_sequence:BH342781"

244235..245799

/note="wgs\_end\_extension"

clone\_end:Sp6"

246218..249063

/note="wgs\_end\_extension"

clone\_end:Sp6"

BASE COUNT 67545 a 45424 c 45961 g 78594 t 13980 others

ORIGIN

Query Match 27.2%; Score 34.8; DB 2; Length 251504;

Best Local Similarity 60.0%; Pred. No. 8.5;

Matches 54; Conservative 2; Mismatches 34; Indels 0; Gaps 0;

QY 8 TACATACACACACTTGTAGTCAGATACCTTTGAAACCTGTTCAAAACCTTAATGCTT 67

Db 87231 TGCAGATGAAAACCTGAGGCACAGGGAACCTCAATACCTTCTCAAGACCATTCGCTT 87172

QY 68 ATAAGARRCTTCAGACAGACAGTCGTGCTC 97

Db 87171 ATAATGAGATTTAACACAGATTTTTTTC 87142

RESULT 12

HSY313F4\_1

WPCOMMENT

Sequence split into 4 fragments

Fragment Name Begin End

Accession AL023808

HSY313P4\_0 1 110000  
 HSY313P4\_1 100001 210000  
 HSY313P4\_2 200001 310000  
 HSY313P4\_3 300001 368509  
 Continuation (2 of 4) of HSY313P4 from base 100001 (AL023808 Homo sapiens chromosome 10)

Query Match 26.9%; Score 34.4; DB 2; Length 110000;  
 Best Local Similarity 60.2%; Pred. No. 12;  
 Matches 53; Conservative 2; Mismatches 33; Indels 0; Gaps 0;

QY 32 TACTACTTTGAAACCTGGTTTCAAAACCTTAATGCTTATAGARFCTTGAGAGACAGTCT 91  
 Db 85693 TAAACCAATAAATCTTTTAAACACAAACACAAACAAAGAAATTAGAAATCCACTCT 85752

QY 92 GTGCTCTGAGTCATAGCGAAGCCATCCC 119  
 Db 85753 GTGCAATGATATTAATGGAATCCCTCCC 85780

RESULT 13  
 ALS911132  
 LOCUS  
 DEFINITION Homo sapiens chromosome 10 clone RP13-260019, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 8 unordered pieces.  
 ACCESSION ALS911132  
 VERSION ALS911132.1 GI:13990328  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1  
 Burton, J.  
 Direct Submission  
 Submitted (18-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: humquery@sanger.ac.uk  
 ----- Project Information  
 Center project name: bB260019  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 157088 bases at least Q40  
 Consensus quality: 157552 bases at least Q30  
 Consensus quality: 157835 bases at least Q20  
 Insert size: 158106; sum-of-contigs  
 Insert size: 165156; 2.3% error; agarose-fp  
 Quality coverage: 7.04x in Q20 bases; sum-of-contigs Quality  
 coverage: 6.79x in Q20 bases; agarose-fp  
 -----

\* NOTE: this is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1  
 \* 22768: contig of 22768 bp in length  
 \* 22769 22868: gap of 100 bp  
 \* 22869 26417: contig of 3549 bp in length  
 \* 26418 26517: gap of 100 bp  
 \* 26518 30459: contig of 3942 bp in length  
 \* 30460 30559: gap of 100 bp  
 \* 30560 58200: contig of 27641 bp in length  
 \* 58201 58300: gap of 100 bp  
 \* 58301 71315: contig of 13015 bp in length



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 14, 2003, 14:14:58 ; Search time 177 Seconds  
(without alignments)  
1952.136 Million cell updates/sec

Title: US-09-889-611-1  
Perfect score: 128  
Sequence: 1 aatgaatacatacaacca.....gaagccatccagaagccag 128

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127.2	99.4	128	21	AAK71434 Human megsin promo
2	127.2	99.4	1431	21	AAK71435 Human megsin promo
c 3	85.4	66.7	6182	24	ABL34015 Human immune syste
4	78.4	61.3	6182	24	ABL34014 Human immune syste
5	32.8	25.6	5363	22	AAK5623 Human immune/haema
6	31.8	24.8	6013	24	AAK63315 Chemically pretrea
c 7	30.8	24.1	12321	25	ABZ72006 Human FUS/TLS prot
8	30.2	23.6	596	22	ABA31153 Probe #9619 for ge

9	30.2	23.6	596	22	AAK12479 Human brain expres
10	30.2	23.6	596	22	AAK38189 Human bone marrow
11	30.2	23.6	596	24	ABS12223 Human genome-deriv
12	30	23.4	30	21	AAK71448 Human megsin promo
13	30	23.4	30	21	AAK71449 Human megsin promo
14	30	23.4	30	21	AAK71450 Human megsin promo
c 15	30	23.4	1394	22	ABA15779 Human nervous syst
c 16	30	23.4	1394	22	AAI62830 Human genomic DNA
17	29.8	23.3	1410	19	AAV53490 DNA encoding a Sta
c 18	29.4	23.0	4671	24	AAK83965 Human gene sequenc
c 19	29.4	23.0	75384	22	AAK85590 Human immune/haema
20	29.2	22.8	82	22	ABA40731 Probe #19197 for g
21	29.2	22.8	82	22	AAK24845 Human brain expres
22	29.2	22.8	82	22	AAK50838 Human bone marrow
23	29.2	22.8	82	24	ABS24338 Human genome-deriv
c 24	29.2	22.8	202001	24	ABS52506 Human transporter
c 25	29	22.7	13063	23	ABU05030 Drosophila melanog
c 26	29	22.7	85680	21	AAF22299 BAC containing rep
c 27	29	22.7	611590	21	AAF22303 Arabidopsis thalia
c 28	28.6	22.3	271	21	AAK05181 Human secreted pro
c 29	28.4	22.2	384	22	AAI87829 Human polynucleoti
30	28.4	22.2	524	22	AAH33044 Human colon cancer
c 31	28.4	22.2	540	22	AAK62913 Human foetal liver
c 32	28.4	22.2	540	22	ABA30183 Probe #8649 for ge
c 33	28.4	22.2	540	22	AAK11321 Human brain expres
c 34	28.4	22.2	540	22	AAK37110 Human bone marrow
c 35	28.4	22.2	540	22	AAI17948 Probe #7881 for ge
c 36	28.4	22.2	540	22	AAI42936 Probe #11622 used
c 37	28.4	22.2	540	23	ABS6791 Human liver single
c 38	28.4	22.2	540	24	ABS11108 Human genome-deriv
39	28.4	22.2	1229	23	ABV24691 Human prostate exp
c 40	28.4	22.2	2219	22	ABA17021 Human nervous syst
c 41	28.4	22.2	3587	25	ACA04011 cDNA downregulated
c 42	28.2	22.0	440	24	ABL84068 Human ovarian canc
c 43	28.2	22.0	5270	24	ABL33039 Human immune syste
c 44	28.2	22.0	6209	22	AAK28751 Genomic sequence #
c 45	28.2	22.0	38342	22	AAK46745 Tumour suppressor

ALIGNMENTS

RESULT 1  
AAK71434  
ID AAK71434 standard; DNA; 128 BP.  
AC AAK71434;  
XX  
XX  
DT 01-DEC-2000 (first entry)  
XX  
Human megsin promoter fragment DNA.  
XX  
Promoter; megsin; human; protein isolation; screening. ss.  
XX  
Homo sapiens.  
XX  
FN WO200043528-A1.  
XX  
PD 27-JUL-2000.  
XX  
PF 25-JAN-2000; 2000WO-JP00350.  
XX  
PR 25-JAN-1999; 95JP-0015667.  
XX  
XX (KURO/) KUROKAWA K.  
XX (MIYA/) MIYATA T.  
XX Miyata T;  
XX WPI; 2000-543257/49.  
XX DNA for promoter region of megsin useful for screening proteins -  
XX

PS Claim 1; Page 32; 45pp; Japanese.  
 CC This invention describes a novel DNA sequence (1) representing a promoter  
 CC region having part or all of a specific base sequence. The invention also  
 CC describes (1) a vector containing (1); (2) a cell transformed by the  
 CC above vector; and (3) protein produced using (1). (1) is useful for  
 CC screening and isolating proteins (especially transcription factors). This  
 CC sequence represents the human megilin promoter which is described in the  
 CC method of the invention.  
 SQ Sequence 128 BP; 44 A; 29 C; 24 G; 29 T; 2 other;  
 Query Match 99.4%; Score 127.2; DB 21; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-31;  
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATGAACATACATACACACACCTTAGTCAGATACATCTTTGAAACCTGGTTCAAAACCTTA 60  
 Db 1 AATGAACATACATACACACACCTTAGTCAGATACATCTTTGAAACCTGGTTCAAAACCTTA 60  
 QY 61 AATGCTTATAGARRCTTGAGAGACAGTCTGTCTGAGTCATAGGAGCCATCCCA 120  
 Db 61 AATGCTTATAGARRCTTGAGAGACAGTCTGTCTGAGTCATAGGAGCCATCCCA 120  
 QY 121 GAAGCCAG 128  
 Db 121 GAAGCCAG 128  
 RESULT 2  
 AAA71435  
 ID AAA71435 standard; DNA; 1431 BP.  
 XX AC AAA71435;  
 XX 01-DEC-2000 (first entry)  
 XX Human megilin promoter fragment DNA.  
 XX Promoter; megilin; human; protein isolation; screening. ss.  
 XX Homo sapiens.  
 XX WO200043528-A1.  
 XX 27-JUL-2000.  
 XX 25-JAN-2000; 2000WO-JP00350.  
 XX 25-JAN-1999; 99JP-0015667.  
 XX (KURO/) KUROKAWA K.  
 XX (MIYA/) MIYATA T.  
 XX Miyata T;  
 XX WPI; 2000-543257/49.  
 XX DNA for promoter region of megilin useful for screening proteins -  
 XX Disclosure; Fig 2; 45pp; Japanese.  
 CC This invention describes a novel DNA sequence (1) representing a promoter  
 CC region having part or all of a specific base sequence. The invention also  
 CC describes (1) a vector containing (1); (2) a cell transformed by the  
 CC above vector; and (3) protein produced using (1). (1) is useful for  
 CC screening and isolating proteins (especially transcription factors). This  
 CC sequence represents a fragment of the human megilin promoter which is  
 CC described in the method of the invention.  
 SQ Sequence 1431 BP; 466 A; 266 C; 303 G; 394 T; 2 other;  
 Query Match 99.4%; Score 127.2; DB 21; Length 1431;

Best Local Similarity 100.0%; Pred. No. 2.6e-31;  
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AATGAACATACATACACACCTTAGTCAGATACATCTTTGAAACCTGGTTCAAAACCTTA 60  
 Db 1304 AATGAACATACATACACACCTTAGTCAGATACATCTTTGAAACCTGGTTCAAAACCTTA 1363  
 QY 61 AATGCTTATAGARRCTTGAGAGACAGTCTGTCTGAGTCATAGGAGCCATCCCA 120  
 Db 1364 AATGCTTATAGARRCTTGAGAGACAGTCTGTCTGAGTCATAGGAGCCATCCCA 1423  
 QY 121 GAAGCCAG 128  
 Db 1424 GAAGCCAG 1431  
 RESULT 3  
 ABL34015/c  
 ID ABL34015 standard; DNA; 6182 BP.  
 XX AC ABL34015;  
 XX 26-MAR-2002 (first entry)  
 XX Human immune system associated gene SEQ ID NO: 1988.  
 XX Human; immune system disease; cytosine methylation; antiasthmatic;  
 KW antiarteriosclerotic; antianaemic; cytostatic; neurotic;  
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
 KW Gene; ds.  
 XX Homo sapiens.  
 XX WO200200928-A2.  
 XX 03-JAN-2002.  
 XX 02-JUL-2001; 2001WO-EP07537.  
 XX 30-JUN-2000; 2000DE-1032529.  
 XX 01-SEP-2000; 2000DE-1043826.  
 XX (EPIC-) EPICENOMICS AG.  
 XX Olek A, Piepenbrock C, Berlin K;  
 XX WPI; 2002-130909/17.  
 XX Nucleic acid comprising fragment of chemically modified gene, useful  
 XX for diagnosis and treatment of diseases associated with abnormal  
 XX cytosine methylation -  
 XX Claim 1; SEQ ID NO 1988; 32pp + Sequence Listing; German.  
 CC The present invention provides a number of human immune system associated  
 CC genes which are modified by the methylation of cytosines. The sequences  
 CC can be used in the diagnosis and treatment of immune system disorders,  
 CC including eye diseases such as retinopathy, neovascular glaucoma and  
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
 CC diseases. The present sequence is a gene of the invention.  
 XX Sequence 6182 BP; 1934 A; 37 C; 1059 G; 3152 T; 0 other;  
 Query Match 66.7%; Score 85.4; DB 24; Length 6182;  
 Best Local Similarity 79.5%; Pred. No. 1.5e-17;  
 Matches 101; Conservative 0; Mismatches 26; Indels 0; Gaps 0;









RESULT 8  
ABA31153  
ID ABA31153 standard; DNA; 596 BP.  
XX  
AC ABA31153;  
XX  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Probe #9619 for gene expression analysis in human heart cell sample.  
XX  
KW Human; gene expression; heart; microarray; vascular system; probe;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157274-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00666.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488899/53.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human hearts -  
XX  
XX  
PS Claim 1; SEQ ID No 9619; 530pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 596 BP; 174 A; 87 C; 162 G; 173 T; 0 other;  
  
Query Match 23.6%; Score 30.2; DB 22; Length 596;  
Best Local Similarity 57.5%; Pred. No. 5.6; Mismatches 35; Indels 0; Gaps 0;  
Matches 50; Conservative 2;  
  
QY 38 TTTCGAACCTGGTTCACAACTTAATGCTTATAGARRCTTGAGAGACAGTGTGCTC 97  
Db 277 TTTCGAAGTGGTTGCATCAACGAGATTTATGAATTTAATACCGAGGATGGT 336  
  
QY 98 TGAGTCATAGGAGGAGCCATCCAGAG 124  
Db 337 GGAGTGTAGGAGGAGCCATCCAGAG 363  
  
RESULT 9  
AAK12479  
ID AAK12479 standard; DNA; 596 BP.  
XX  
AC AAK12479;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe SEQ ID NO: 12470.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human brains -  
XX  
XX  
PS Example 4; SEQ ID NO: 12470; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.  
XX  
SQ Sequence 596 BP; 174 A; 87 C; 162 G; 173 T; 0 other;  
  
Query Match 23.6%; Score 30.2; DB 22; Length 596;  
Best Local Similarity 57.5%; Pred. No. 5.6; Mismatches 35; Indels 0; Gaps 0;  
Matches 50; Conservative 2;  
  
QY 38 TTTCGAACCTGGTTCACAACTTAATGCTTATAGARRCTTGAGAGACAGTGTGCTC 97  
Db 277 TTTCGAAGTGGTTGCATCAACGAGATTTATGAATTTAATACCGAGGATGGT 336  
  
QY 98 TGAGTCATAGGAGGAGCCATCCAGAG 124  
Db 337 GGAGTGTAGGAGGAGCCATCCAGAG 363  
  
RESULT 10  
AAK38189  
ID AAK38189 standard; DNA; 596 BP.  
XX  
AC AAK38189;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed single exon probe SEQ ID NO: 12746.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

OS WO200157276-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 12746; 658pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is one of

CC the probes of the invention.

CC Sequence 596 BP; 174 A; 87 C; 162 G; 173 T; 0 other;

CC Query Match 23.6%; Score 30.2; DB 22; Length 596;

CC Best Local Similarity 57.5%; Pred. No. 5.6;

CC Matches 50; Conservative 2; Mismatches 35; Indels 0; Gaps 0;

QY 38 TTGTAAACCTGTTTCAAACTTAATCTTATAGARCTTGAGAGACAGTGTGCTC 97

DB 277 TTGTAGAGATGTTGTCATGCAAAACGAAGTATTATGAAATTTTAATACCGAGGAATGGCT 336

QY 98 TGAGTCATAGGAGGACGCTCCAGAG 124

DB 337 GGAGTGTAGGAGGACGCTCCAGAG 363

RESULT 11

ABSL12223

ID ABSL12223 standard; DNA; 596 BP.

XX ABSL12223;

DT 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe from lung SEQ ID No 12214.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

PN

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

WO200186003-A2.

15-NOV-2001.

30-JAN-2001; 2001WO-US00665.

04-FEB-2000; 2000US-180312P.

26-MAY-2000; 2000US-207456P.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-234687P.

27-SEP-2000; 2000US-236359P.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to

measure gene expression in human lung samples -

Claim 1; SEQ ID No 12214; 634pp; English.

The invention relates to a spatially-addressable set of single exon

nucleic acid probes for measuring gene expression in a sample derived

from human lung comprising single exon nucleic acid probes having one of

12614 nucleic acid sequences mentioned in the specification, or their

complements or the 12387 open reading frames derived from the 12614

probes. Also included are a microarray comprising the novel set of

probes; the novel set of probes which hybridise at high stringency to a

nucleic acid expressed in the human lung; measuring gene expression in a

sample derived from human lung, comprising (a) contacting the array with

a collection of detectably labeled nucleic acids derived from human lung

mRNA, and (b) measuring the label detectably bound to each probe of

the array; identifying exons in a eukaryotic genome, comprising

(a) algorithmically predicting at least one exon from genomic sequences

of the eukaryote; and (b) detecting specific hybridisation of detectably

labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

having a fragment identical to the predicted exon, the probe is included

in the above mentioned microarray; assigning exons to a single gene,

comprising (a) identifying exons from genomic sequence by the method

above and (b) measuring the expression of each of the exons in several

tissues and/or cell types using hybridisation to a single exon

microarrays having a probe with the exon, where a common pattern of

expression of the exons in the tissues and/or cell types indicates that

the exons should be assigned to a single gene; a peptide comprising one

of 12011 sequences, mentioned in the specification, or encoded by the

probes/open reading frames (ORF). The probes are used for gene

expression analysis, and for identifying exons in a gene, particularly

using human lung derived mRNA and for the study of lung diseases

such as asthma, lung cancer, chronic obstructive pulmonary disease

(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,

pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic

pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension

and hyaline membrane disease. The present sequence is a single exon

probe of the invention.

Note: The sequence data for this patent did not form part

of the printed specification, but was obtained in electronic

format directly from WIPO at

ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 596 BP; 174 A; 87 C; 162 G; 173 T; 0 other;

Query Match 23.6%; Score 30.2; DB 24; Length 596;

Best Local Similarity 57.5%; Pred. No. 5.6;

Matches 50; Conservative 2; Mismatches 35; Indels 0; Gaps 0;

QY 38 TTGTAAACCTGTTTCAAACTTAATCTTATAGARCTTGAGAGACAGTGTGCTC 97

DB 277 TTGTAGAGATGTTGTCATGCAAAACGAAGTATTATGAAATTTTAATACCGAGGAATGGCT 336

QY 98 TGAGTCATAGGAGGACGCTCCAGAG 124

DB 337 GGAGTGTAGGAGGACGCTCCAGAG 363

RESULT 11

ABSL12223

ID ABSL12223 standard; DNA; 596 BP.

XX ABSL12223;

DT 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe from lung SEQ ID No 12214.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

XX familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

XX Homo sapiens.



CC promoter region having part or all of a specific base sequence. The  
CC invention also describes (1) a vector containing (1); (2) a cell  
CC transformed by the above vector; and (3) protein produced using (1). (1)  
CC is useful for screening and isolating proteins (especially transcription  
CC factors). AAT71434-A71469 represent PCR primers used in the method  
CC described in the invention.

XX  
SQ Sequence 30 BP; 11 A; 7 C; 3 G; 9 T; 0 other;

Query Match 23.4%; Score 30; DB 21; Length 30;

Best Local Similarity 100.0%; Pred. No. 2.6; Mismatches 0; Indels 0; Gaps 0;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 TACTACTTTGAAACCTGGTTCAAAACCTAA 61

Db 1 TACTACTTTGAAACCTGGTTCAAAACCTAA 30

RESULT 15

ABAL5779/c

ID ABAL5779 standard; DNA; 1394 BP.

XX

AC ABAL5779;

XX

DT 23-JAN-2002 (first entry)

XX

DE Human nervous system related polynucleotide SEQ ID NO 8110.

XX

KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX

OS Homo sapiens.

XX

PN WO200159063-A2.

XX

PD 16-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01334.

XX

PR 31-JAN-2000; 2000US-0179065.

PR

PR 04-FEB-2000; 2000US-0180628.

PR

PR 24-FEB-2000; 2000US-0184664.

PR

PR 02-MAR-2000; 2000US-0186350.

PR

PR 16-MAR-2000; 2000US-0189874.

PR

PR 17-MAR-2000; 2000US-0190076.

PR

PR 18-APR-2000; 2000US-0198123.

PR

PR 19-MAY-2000; 2000US-0205515.

PR

PR 07-JUN-2000; 2000US-0209467.

PR

PR 28-JUN-2000; 2000US-0214886.

PR

PR 30-JUN-2000; 2000US-0215135.

PR

PR 07-JUL-2000; 2000US-0216647.

PR

PR 07-JUL-2000; 2000US-0216880.

PR

PR 11-JUL-2000; 2000US-0217487.

PR

PR 11-JUL-2000; 2000US-0217496.

PR

PR 14-JUL-2000; 2000US-0218290.

PR

PR 26-JUL-2000; 2000US-0220963.

PR

PR 14-AUG-2000; 2000US-0220964.

PR

PR 14-AUG-2000; 2000US-0224518.

PR

PR 14-AUG-2000; 2000US-0224519.

PR

PR 14-AUG-2000; 2000US-0225213.

PR

PR 14-AUG-2000; 2000US-0225214.

PR

PR 14-AUG-2000; 2000US-0225266.

PR

PR 14-AUG-2000; 2000US-0225267.

PR

PR 14-AUG-2000; 2000US-0225268.

PR

PR 14-AUG-2000; 2000US-0225270.

PR

PR 14-AUG-2000; 2000US-0225447.

PR

PR 14-AUG-2000; 2000US-0225757.

PR

PR 14-AUG-2000; 2000US-0225758.

PR

PR 14-AUG-2000; 2000US-0225759.

PR

PR 14-AUG-2000; 2000US-0225760.

PR

PR 14-AUG-2000; 2000US-0225761.

PR

PR 14-AUG-2000; 2000US-0225762.

PR

PR 14-AUG-2000; 2000US-0225763.

PR

PR 14-AUG-2000; 2000US-0225764.

PR

PR 14-AUG-2000; 2000US-0225765.

PR

PR 14-AUG-2000; 2000US-0225766.

PR

PR 14-AUG-2000; 2000US-0225767.

PR

PR 14-AUG-2000; 2000US-0225768.

PR

PR 14-AUG-2000; 2000US-0225769.

PR

PR 14-AUG-2000; 2000US-0225770.

PR

PR 14-AUG-2000; 2000US-0225771.

PR

PR 14-AUG-2000; 2000US-0225772.

PR

PR 14-AUG-2000; 2000US-0225773.

PR

PR 14-AUG-2000; 2000US-0225774.

PR

PR 14-AUG-2000; 2000US-0225775.

PR

PR 14-AUG-2000; 2000US-0225776.

PR

PR 14-AUG-2000; 2000US-0225777.

PR

PR 14-AUG-2000; 2000US-0225778.

PR

PR 14-AUG-2000; 2000US-0225779.

PR

PR 14-AUG-2000; 2000US-0225780.

PR

PR 14-AUG-2000; 2000US-0225781.

PR

PR 14-AUG-2000; 2000US-0225782.

PR

PR 14-AUG-2000; 2000US-0225783.

PR

PR 14-AUG-2000; 2000US-0225784.

PR

PR 14-AUG-2000; 2000US-0225785.

PR

PR 14-AUG-2000; 2000US-0225786.

PR

PR 14-AUG-2000; 2000US-0225787.

PR

PR 14-AUG-2000; 2000US-0225788.

PR

PR 14-AUG-2000; 2000US-0225789.

PR

PR 14-AUG-2000; 2000US-0225790.

PR

PR 14-AUG-2000; 2000US-0225791.

PR

PR 14-AUG-2000; 2000US-0225792.

PR

PR 14-AUG-2000; 2000US-0225793.

PR

PR 14-AUG-2000; 2000US-0225794.

PR

PR 14-AUG-2000; 2000US-0225795.

PR

PR 14-AUG-2000; 2000US-0225796.

PR

PR 14-AUG-2000; 2000US-0225797.

PR

PR 14-AUG-2000; 2000US-0225798.

PR

PR 14-AUG-2000; 2000US-0225799.

PR

PR 14-AUG-2000; 2000US-0225800.

PR

PR 14-AUG-2000; 2000US-0225801.

PR

PR 14-AUG-2000; 2000US-0225802.

PR

PR 14-AUG-2000; 2000US-0225803.

PR

PR 14-AUG-2000; 2000US-0225804.

PR

PR 14-AUG-2000; 2000US-0225805.

PR

PR 14-AUG-2000; 2000US-0225806.

PR

PR 14-AUG-2000; 2000US-0225807.

PR

PR 14-AUG-2000; 2000US-0225808.

PR

PR 14-AUG-2000; 2000US-0225809.

PR

PR 14-AUG-2000; 2000US-0225810.

PR

PR 14-AUG-2000; 2000US-0225811.

PR

PR 14-AUG-2000; 2000US-0225812.

PR

PR 14-AUG-2000; 2000US-0225813.

PR

PR 14-AUG-2000; 2000US-0225814.

PR

PR 14-AUG-2000; 2000US-0225815.

PR

PR 14-AUG-2000; 2000US-0225816.

PR

PR 14-AUG-2000; 2000US-0225817.

PR

PR 14-AUG-2000; 2000US-0225818.

PR

PR 14-AUG-2000; 2000US-0225819.

PR

PR 14-AUG-2000; 2000US-0225820.

PR

PR 14-AUG-2000; 2000US-0225821.

PR

PR 14-AUG-2000; 2000US-0225822.

PR

PR 14-AUG-2000; 2000US-0225823.

PR

PR 14-AUG-2000; 2000US-0225824.

PR

PR 14-AUG-2000; 2000US-0225825.

PR

PR 14-AUG-2000; 2000US-0225826.

PR

PR 14-AUG-2000; 2000US-0225827.

PR

PR 14-AUG-2000; 2000US-0225828.

PR

PR 14-AUG-2000; 2000US-0225829.

PR

PR 14-AUG-2000; 2000US-0225830.

PR

PR 14-AUG-2000; 2000US-0225831.

PR

PR 14-AUG-2000; 2000US-0225832.

PR

PR 14-AUG-2000; 2000US-0225833.

PR

PR 14-AUG-2000; 2000US-0225834.

PR

PR 14-AUG-2000; 2000US-0225835.

PR

PR 14-AUG-2000; 2000US-0225836.

PR

PR 14-AUG-2000; 2000US-0225837.

PR

PR 14-AUG-2000; 2000US-0225838.

PR

PR 14-AUG-2000; 2000US-0225839.

PR

PR 14-AUG-2000; 2000US-0225840.

PR

PR 14-AUG-2000; 2000US-0225841.

PR

PR 14-AUG-2000; 2000US-0225842.

PR

PR 14-AUG-2000; 2000US-0225843.

PR

PR 14-AUG-2000; 20

PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX WPI; 2001-541565/60.  
XX  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating nervous system  
PT cancers and metastases -  
XX  
XX Disclosure; SEQ ID NO 8110; 1701pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (AB11004-ABA21534) and proteins  
CC (AB114678-AB18001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 1394 BP; 313 A; 385 C; 338 G; 358 T; 0 other;  
SQ  
  
Query Match 23.4%; Score 30; DB 22; Length 1394;  
Best Local Similarity 55.1%; Pred. NO. 8.5;  
Matches 54; Conservative 2; Mismatches 42; Indels 0; Gaps 0;  
  
QY 4 GAACCTACATACACCACTTAGTCAGATACCTCTTGAAACCTGGTTCAAAACCTAAAT 63

Db 811 GAATTTTGTATTAGCACTCAATAAATACATATATCCCCCACCTTCAAGCCTAGAA 752  
QY 64 GTTATAAGARRCTTGGAGACAGACGTGCTGTGCTGAG 101  
Db 751 GTTGACTTGAGGCCTTAGAAACAATTTATCTGCAATCAG 714

Search completed: August 14, 2003, 17:00:10  
Job time : 180 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 16:46:23 ; Search time 1667 Seconds  
(without alignments)  
1866.211 Million cell updates/sec

Title: US-09-889-611-1

Perfect score: 128

Sequence: 1 aatgaactacataacaacca.....gaagccatccagaagcag 128

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	34.8	27.2	575	28	AQ993609
C 2	33.6	26.2	559	28	BH006869
C 3	33.2	25.9	671	13	EX250069
C 4	32.6	25.5	613	9	AW695053
					AW695053 NF082H10S

# ALIGNMENTS

RESULT 1	AQ993609/c	575 bp	DNA	linear	GSS 24-FEB-2000
LOCUS	RPCI-23-270J20.TV	RPCI-23	Mus musculus	genomic clone	RPCI-23-270J20
DEFINITION	genomic survey sequence.				
ACCESSION	AQ993609	GI:7068706			
VERSION	AQ993609.1	GI:7068706			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 575)				
AUTHORS	Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret ,B., Levine,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.				
TITLE	Mouse BAC End Sequences from Library RPCI-23				
JOURNAL	Unpublished				
COMMENT	Other_GSSs: RPCI-23-270J20.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhao@tigr.org Clones are derived from the mouse BAC library RPCI-23. For BAC				

BZ263242 CH230-521  
AQ377684 RPT111-15  
AA321831 RST24393  
BJ378830 BJ378830  
BQ699205 NXRV120 G  
AZ145433 SP\_0021\_B  
AW225931 ST75D05 F  
BQ655581 NXRV096 D  
BQ699018 NXRV118 D  
BE431403 NXRV 181  
BQ655089 NXRV090 B  
BQ700311 NXRV103 H  
CC322562 TAM32-34C  
BX381001 BX381001  
AA627611 nq46h03. s  
BW300166 BW300166  
BY675226 BY675226  
AV850545 AV850545  
AV841822 AV841822  
BW295254 BW295254  
AQ795888 nbxb0058B  
BW010244 BW010244  
BW300165 BW300165  
BW025989 BW025989  
BW165883 BW165883  
BW170317 BW170317  
BH519876 BOHMC82TF  
D55583 HUM183B04B  
AQ838161 HS\_4712 A  
CC059231 ii21a11.b  
BU769759 SUECIE08  
BZ780856 ii21a11.g  
BZ876431 CH240 214  
AW042772 ST25C01 P  
BH972288 odg72509.  
AQ321395 RPCI11-10  
BQ699354 NXRV126 F  
BJ508716 BJ508716  
BJ503935 BJ503935  
CD358891 AGENCOURT  
BF166530 601776529

library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 270 row: J column: 20  
 Seq primer: SP6  
 Class: BAC ends.

#### FEATURES

Location/Qualifiers

1..575  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-270J20"  
 /sex="female"  
 /lab\_host="DH10B"  
 /clone\_lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Site selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
 " "

BASE COUNT 153 a 93 c 119 g 209 t 1 others

#### ORIGIN

Query Match 27.2%; Score 34.8; DB 28; Length 575;  
 Best Local Similarity 78.0%; Pred. No. 8.9;  
 Matches 39; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 52 CAAAACCTAATGCTTATAGACRCTTGACAGACAGCTGCTGCTCAG 101  
 DB 90 CAAAACCTAATGCTTATAGACRCTTGACAGACAGCTGCTGCTCAG 41

#### RESULT 2

BH006869  
 LOCUS BH006869  
 DEFINITION BHEACIN1977 PSU Brugia malayi genomic BAC Library 1 & 2 Brugia malayi genomic, genomic survey sequence.

ACCESSION BH006869.2 GI:19585607  
 KEYWORDS GSS.

#### SOURCE

Brugia malayi  
 Brugia malayi  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Brugia.

REFERENCE 1 (bases 1 to 559)  
 Whitton,C., Daub,J., Ware,J., Quail,M., Hall,N., Barrell,B., Foster,J., Gulliano,D., Slatko,B. and Blaxter,M.

TITLE Genome survey sequences from the human parasitic nematode Brugia malayi

#### JOURNAL

Unpublished

On Mar 21, 2002 this sequence version replaced gi:13949922.

Contact: Blaxter ML

Institute of Cell, Animal and Population Biology

University of Edinburgh

Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK

Tel: +44 131 650 6760

Fax: +44 131 670 5450

Email: mark.blaxter@ed.ac.uk

Sequenced from the Filarial Genome Project's Brugia malayi BAC library constructed by Jesse Pope-Chappel and Jeremy Foster. The sequence was generated by The Pathogen Sequencing Unit, The Sanger Centre, Cambridge, UK in collaboration with Mark Blaxter, ICAPB, University of Edinburgh, Edinburgh, UK.

Seq primer: T7 (TAATACGACTCACTATAGGG)

Class: BAC ends.

Location/Qualifiers

1..559

#### FEATURES

source

/organism="Brugia malayi"  
 /mol\_type="genomic DNA"  
 /strain="TRS"  
 /db\_xref="taxon:6279"  
 /sex="Mixed (male and female)"  
 /tissue\_type="whole parasite"  
 /dev stage="adult"

/clone\_lib="Brugia malayi Genomic BAC Library 1 & 2"

/note="Vector: pBelOBAC II; Site 1: Hind III; Brugia malayi genomic DNA was partially cleaved with Hind III and size fractionated. 18,000 clones were generated from 2 libraries with mean insert size 60 kbp. The library was constructed by Jesse Pope-Chappel, Smith College Northampton MA and Dr Jeremy Foster, New England Biolabs, MA."

BASE COUNT 232 a 84 c 84 g 159 t

#### ORIGIN

Query Match 26.2%; Score 33.6; DB 28; Length 559;

Best Local Similarity 66.7%; Pred. No. 20;

Matches 48; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 2 ATGAACCTACATACACACACCTTAGTCAGATACCTTTGAAACCTGGTTCAAAACCTAA 61

DB 7 AUGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 66

QY 62 ATGCTTATAAGA 73

DB 67 TTCTTGATAATA 78

#### RESULT 3

BX250069/c

LOCUS BX250069

DEFINITION Pinus pinaster differentiating xylem adult Pinus pinaster

sequence.

ACCESSION BX250069.1 GI:28510204

VERSION BX250069

KEYWORDS EST.

SOURCE Pinus pinaster

ORGANISM Pinus pinaster

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus;

1 (bases 1 to 671)

REFERENCE 1 (bases 1 to 671)

AUTHORS Canton,F.R., Le Provost,G., Garcia,V., Barre,A., Frigerio,J.-M.,

Paiva,J., Feveiro,P., Avila,C., Mouret,J.-F., Brach,J., de

Daruvur,A., Canovas,F.M. and Plomion,C.

Transcriptome analysis of wood formation in maritime pine

Unpublished

Contact: Frigerio JM

Genetique et Amelioration 69

INRA

route d'Arcachon 33612 Cestas CEDEX France

Email: Frigerio@pierrot.inra.fr

Email: Frigerio@pierrot.inra.fr.

Location/Qualifiers

1..671

/organism="Pinus pinaster"

/mol\_type="mRNA"

/strain="ecotype: Corsican"

/db\_xref="taxon:71647"

/clone="PF032A10"

/tissue\_type="differentiating xylem"

/dev stage="adult"

/clone\_lib="Pinus pinaster differentiating xylem adult"

/note="Vector: Uni-Zap XR lambda (Stratagene); Site 1: Eco

RI; Site 2: Xho I; A composite cDNA library was made with

mRNA isolated from normal, compression, opposite, early

and late wood of Maritime pine uni-directionally cloned

into Uni-Zap XR using the ZAP-CDNA Synthesis kit

(Stratagene). pBluescript SK(-) plasmids were obtained by

in vivo mass excision. The nucleotide sequence of the



```

Db      225 GCAGAACTGAAAAACAGTTTTTCTTAACCACTTTAAGATAAAAATTCCAAGAA 171
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 5
BZ263242
LOCUS           BZ263242                858 bp    DNA        linear   GSS 15-OCT-2002
DEFINITION     CH230-521E18.TJC CHORI-230 Segment 2 Rattus norvegicus genomic
clone CH230-521E18, genomic survey sequence.
ACCESSION      BZ263242
VERSION        BZ263242.1 GI:23974604
KEYWORDS       GSS.
SOURCE         Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE      1 (bases 1 to 858)
AUTHORS        Zhao,S., Shetty,J.S., Shatsman,S., Tsegaye,G., Geer,K., Shivartsbeyn
              A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
              Jong,P. and Fraser,C.W.
TITLE          Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL        Unpublished
COMMENT        Other GSSs: CH230-521E18.TVB
               Contact: Shaying Zhao
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel: 301 838 0200
               Fax: 301 838 0208
               Email: szhao@tigr.org
               Clones are derived from the rat BAC library CHORI-230
               (http://www.chori.org/bacpac/rat230.htm). For BAC library
               availability, please contact Pieter de Jong [pjejong@mail.choi.org].
               Clones may be purchased from BACPAC Resources
               (http://www.chori.org/bacpac/orering\_information.htm). BAC end
               plate: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html
               Page: 521 row: E column: 18
               Seq primer: SP6
               Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..858
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-521E18"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 2"
                     /note=vector: PTABAC1.3; Site 1: MboI; Site 2: MboI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      284 a 178 c 175 g 221 t
ORIGIN
Query Match      25.5%; Score 32.6; DB 29; Length 858;
Best Local Similarity 58.2%; Pred.No.50;
Matches 53; Conservative 2; Mismatches 36; Indels 0; Gaps 0;
QY      1 AATGAACACTACAATAACACCACCTTAGTCAGATPACTACTTGAAACCTGGTCCAAACCTA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      172 AGTGAAGCAGCACCTTCAGTCACTCAGTTACCTATTCCTTTAAATACGTGATTTTAAATC 231
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      61 AATGCTTATATAGARRCTTTGAGACAGTGCT 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      232 AATGCTAATATGCAATTTCAAAGTCAGGGAT 262
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 6
AQ377684/c
LOCUS           AQ377684                698 bp    DNA        linear   GSS 20-MAY-1999
DEFINITION     RPC111-151114.TU RPC1-11 Homo sapiens genomic clone RPC1-11-151114
clone RP11-151114, genomic survey sequence.
ACCESSION      AQ377684
VERSION        AQ377684.1 GI:23974604
KEYWORDS       GSS.
SOURCE         Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE      1 (bases 1 to 858)
AUTHORS        Zhao,S., Shetty,J.S., Shatsman,S., Tsegaye,G., Geer,K., Shivartsbeyn
              A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
              Jong,P. and Fraser,C.W.
TITLE          Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL        Unpublished
COMMENT        Other GSSs: CH230-521E18.TVB
               Contact: Shaying Zhao
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel: 301 838 0200
               Fax: 301 838 0208
               Email: szhao@tigr.org
               Clones are derived from the rat BAC library CHORI-230
               (http://www.chori.org/bacpac/rat230.htm). For BAC library
               availability, please contact Pieter de Jong [pjejong@mail.choi.org].
               Clones may be purchased from BACPAC Resources
               (http://www.chori.org/bacpac/orering\_information.htm). BAC end
               plate: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html
               Page: 521 row: E column: 18
               Seq primer: SP6
               Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..858
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-521E18"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 2"
                     /note=vector: PTABAC1.3; Site 1: MboI; Site 2: MboI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
BASE COUNT      284 a 178 c 175 g 221 t
ORIGIN
Query Match      25.5%; Score 32.6; DB 29; Length 858;
Best Local Similarity 58.2%; Pred.No.50;
Matches 53; Conservative 2; Mismatches 36; Indels 0; Gaps 0;
QY      1 AATGAACACTACAATAACACCACCTTAGTCAGATPACTACTTGAAACCTGGTCCAAACCTA 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      172 AGTGAAGCAGCACCTTCAGTCACTCAGTTACCTATTCCTTTAAATACGTGATTTTAAATC 231
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      61 AATGCTTATATAGARRCTTTGAGACAGTGCT 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      232 AATGCTAATATGCAATTTCAAAGTCAGGGAT 262
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

/dev stage="Culmination stage"
/clone lib="Dictyostellium discoideum cDNA library, CF"
BASE COUNT 263 a 125 c 146 g 228 t 2 others
ORIGIN
Query Match 25.2%; Score 32.2; DB 12; Length 764;
Best Local Similarity 70.5%; Pred. No. 62;
Matches 43; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 AATGAACACTACATACACACCTTACTGATACATCTTGAACACCTGTTCAAAACCTA 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
255 ATTGAACCTGCAACATTCACCTGGAGCAGTAAGGCTTTGTAACTGGTTCAACACCTA 314
QY 61 A 61
Db 315 A 315
RESULT 9
BQ699205/c
LOCUS
DEFINITION BQ699205 386 bp mRNA linear EST 07-MAY-2003
ACCESSION BQ699205
VERSION BQ699205.1 GI:21824521
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
TITLE 1 (bases 1 to 386)
JOURNAL Sederoff, R.
COMMENT Molecular Basis of Wood Formation in the Pine Megagenome
Contact: Sederoff, Ron
Forest Biotechnology
North Carolina State University
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
NC 27695, USA
Tel: 919 515 7800
Fax: 919 515 7801
Email: ron.sederoff@ncsu.edu
Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further
information.
Seq primer: T3.
FEATURES
Location/Qualifiers
source
1..386
/mol_type="mRNA"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXXRV120.G11"
/tissue_type="Xylem"
/cell_type="Root (primary)"
/dev_stage="transitional"
/lab_host="X11-Blue"
/clone_lib="NXXRV (Nsf Xylem Root wood Vertical)"
/notes="Vector: pBlueScript SK-; Site 1: Eco RI; Site 2:
XhoI; The library is from primary xylem scraped from the
roots of a twelve year old tree in the transitional phase
from juvenile wood to mature wood production. NOTE: The
sequences contain a 'cDNA adapter' between the EcoRI site
and the start of the EST. The adapter sequence is
'AATTCGGCAGG'."
BASE COUNT 104 a 60 c 65 g 142 t 15 others
ORIGIN
Query Match 25.0%; Score 32; DB 13; Length 386;
Best Local Similarity 59.5%; Pred. No. 49;
Matches 50; Conservative 2; Mismatches 32; Indels 0; Gaps 0;
QY 1 AATGAACACTACATACACACCTTACTGATACATCTTGAACACCTGTTCAAAACCTA 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 341 AATGAAGTACATACAAAACAATAATTTTCCAAATTTGAAAATTTGATGCAAGAAAT 282
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 AAAGCTCTAAGAACTTGAAGA 258
Db
RESULT 10
AZ145433
LOCUS
DEFINITION AZ145433 528 bp DNA linear GSS 28-AUG-2000
SP_0021_B1_C05_T7 Strongylocentrotus purpuratus, purple sea urchin,
sperm genomic BAC library Strongylocentrotus purpuratus genomic
clone.Plates=21 Col=9 Row=F, genomic survey sequence.
ACCESSION AZ145433
VERSION AZ145433.1 GI:8297336
KEYWORDS GSS.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
REFERENCE Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
AUTHORS Echinoida; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 528)
Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R.,
Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., Wray
G.A., Ettensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H. and
Hood, L.
A sea urchin genome project: Sequence scan, virtual map, and
additional resources
Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
20402566
PUBMED 10920195
COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 21 row: F column: 9
Seq primer: T7
Class: BAC ends
High quality sequence stop: 528.
FEATURES
Location/Qualifiers
source
1..528
/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone="Plate=21 Col=9 Row=F"
/clone_lib="Strongylocentrotus purpuratus, purple sea
urchin, sperm genomic BAC library"
/notes="Organ: sperm; Vector: BAC3.6; BAC Clones in B-Coli
DH10B"
BASE COUNT 161 a 104 c 82 g 177 t 4 others
ORIGIN
Query Match 24.8%; Score 31.8; DB 28; Length 528;
Best Local Similarity 52.0%; Pred. No. 67;
Matches 66; Conservative 2; Mismatches 59; Indels 0; Gaps 0;
QY 1 AATGAACACTACATACACACCTTACTGATACATCTTGAACACCTGTTCAAAACCTA 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
295 AATGAACACTATATAAAACCTGTTTACGCTGATAGTCCATTATTAATGATTAATACATTA 354
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 AATGCTTATAGARRCTTGAGAGACAGCTGCTGCTCTGAGTCATAGGGAACCCATCCA 120
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 GACAGCTATGAGGTACGGTTTGCAATGATGTTAATTAATGTTAAAGTTACCCCA 414
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 GAAGCCA 127
QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 CAGACCA 421
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 11

```

```

AW225931/c
LOCUS       AW225931               339 bp    mRNA    linear    EST 10-DEC-1999
DEFINITION   S775D05 Pine Triplex shoot tip library Pinus taeda cDNA clone
              S775D05, mRNA sequence.
ACCESSION   AW225931
VERSION     AW225931.1  GI:6555227
KEYWORDS    EST.
SOURCE      Pinus taeda (loblolly pine)
            BUKARYOKA; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 339)
AUTHORS     Whetten,R.W., Kinlaw,C.S., Retzel,E. and Sederoff,R.R.
TITLE       The Pine Gene Discovery Project
JOURNAL     Unpublished
COMMENT     Contact: Ross Whetten
            Forest Biotechnology Group
            North Carolina State University
            Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
            ,NC, 27695-8008
            Tel: 919-515-7800
            Fax: 919-515-7801
            Email: rosswhetten@unity.ncsu.edu
            Seq primer: 5' lambda Triplex2 Sequencing Primer.
FEATURES             Location/Qualifiers
     source          1..339
                     /organism="Pinus taeda"
                     /mol_type="mRNA"
                     /db_xref="taxon:3352"
                     /clone="S775D05"
                     /lab_host="E. coli BM25.8"
                     /clone_lib="Pine Triplex shoot tip library"
                     /note="Organ: shoot tips; Vector: Lambda Triplex; Site 1:
                     SfiI (A); Site 2: SfiI (B); Shoot tips (approx. 2 cm from
                     apex) were collected during the spring, frozen and used
                     for mRNA isolation. The SMART-PCR method (Clontech) was
                     used to prepare a library from 1 ug total RNA, using the
                     Lambda Triplex vector. Plasmid subclones in pTriplex were
                     recovered by cre-lox excision in E. coli strain EM25.8 and
                     sequenced from the 5' end."
BASE COUNT      109 a    50 c    79 g    92 t      9 others
ORIGIN
Query Match      24.7%; Score 31.6; DB 9; Length 339;
Best Local Similarity 59.1%; Pred. No. 60;
Matches 52; Conservative 1; Mismatches 35; Indels 0; Gaps 0;

QY  7  CTACATACACACCTTAGTCAGATCTACTTGAACCTGGTTCAAACCTTAATGCT 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  138 CTTCTTAATAACCTCTTTGCAAGTACTTGTCTGTGAATTGCTCACAATCAACAGTGCT 79

QY  67 TATAAGARRCTTGAGACAGCTGCTGTG 94
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  78 TATATCAATCTTGGTGAGGTTGCAATG 51

RESULT 12
LOCUS       B0655581/c              438 bp    mRNA    linear    EST 07-MAY-2003
DEFINITION   NXR096 D12 F NXRV (Nsf Xylem Root wood Vertical) Pinus taeda cDNA
              clone NXR096 D12 5' similar to Arabidopsis thaliana sequence
              At1g18540 unknown protein see
              http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.
ACCESSION   B0655581
VERSION     B0655581.1  GI:21787907
KEYWORDS    EST.
SOURCE      Pinus taeda (loblolly pine)
            BUKARYOKA; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 438)
AUTHORS     Sederoff,R.
TITLE       Molecular Basis of Wood Formation in the Pine Megagenome
JOURNAL     Unpublished
COMMENT     Contact: Sederoff, Ron
            Forest Biotechnology
            North Carolina State University
            840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
            NC 27695, USA
            Tel: 919 515 7800
            Fax: 919 515 7801
            Email: ron_sederoff@ncsu.edu

```

```

Unpublished
Contact: Sederoff, Ron
Forest Biotechnology
North Carolina State University
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
NC 27695, USA
Tel: 919 515 7800
Fax: 919 515 7801
Email: ron_sederoff@ncsu.edu
Please see http://web.ahc.unc.edu/biodata/nsfpine/ for further
information.
Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..438
                     /organism="Pinus taeda"
                     /mol_type="mRNA"
                     /strain="Coastal plain loblolly pine from North Carolina"
                     /db_xref="taxon:3352"
                     /clone="NXRV096 D12"
                     /tissue_type="Xylem"
                     /cell_type="Root (primary)"
                     /dev_stage="Transitional"
                     /lab_host="XLI-Blue"
                     /clone_lib="NXRV (Nsf Xylem Root wood Vertical)"
                     /note="Vector: pBlueScript SK-; Site 1: Eco RI; Site 2:
                     XhoI; The library is from primary xylem scraped from the
                     roots of a twelve year old tree in the transitional phase
                     from juvenile wood to mature wood production. NOTE: The
                     sequences contain a 'cDNA adapter' between the EcoRI site
                     and the start of the EST. The adapter sequence is
                     'AATTCGACACGAG'."
BASE COUNT      122 a    83 c    98 g    121 t    14 others
ORIGIN
Query Match      24.7%; Score 31.6; DB 13; Length 438;
Best Local Similarity 59.1%; Pred. No. 69;
Matches 52; Conservative 1; Mismatches 35; Indels 0; Gaps 0;

QY  7  CTACATACACACCTTAGTCAGATCTACTTGAACCTGGTTCAAACCTTAATGCT 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  266 CTTCTTAATAACCTCTTTGCAAGTACTTGTCTGTGAATTGCTCACAATCAACAGTGCT 207

QY  67 TATAAGARRCTTGAGACAGCTGCTGTG 94
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  206 TATATCAATCTTGGTGAGGTTGCAATG 179

RESULT 13
LOCUS       B0699018/c              507 bp    mRNA    linear    EST 07-MAY-2003
DEFINITION   NXR018 D07 F NXRV (Nsf Xylem Root wood Vertical) Pinus taeda cDNA
              clone NXR018 D07 5' similar to Arabidopsis thaliana sequence
              At1g18540 unknown protein see
              http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.
ACCESSION   B0699018
VERSION     B0699018.1  GI:21824334
KEYWORDS    EST.
SOURCE      Pinus taeda (loblolly pine)
            BUKARYOKA; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 507)
AUTHORS     Sederoff,R.
TITLE       Molecular Basis of Wood Formation in the Pine Megagenome
JOURNAL     Unpublished
COMMENT     Contact: Sederoff, Ron
            Forest Biotechnology
            North Carolina State University
            840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
            NC 27695, USA
            Tel: 919 515 7800
            Fax: 919 515 7801
            Email: ron_sederoff@ncsu.edu

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 17:56:33 ; Search time 1630 Seconds  
(without alignments)  
3212.535 Million cell updates/sec

Title: US-09-889-611-1

Perfect score: 128

Sequence: 1 aatgaactacataacaacaa.....gaagccatccagagccag 128

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_hgt\_hum.\*

31: em\_hgt\_inv.\*

32: em\_hgt\_other.\*

33: em\_hgt\_mus.\*

34: em\_hgt\_pln.\*

35: em\_hgt\_rod.\*

36: em\_hgt\_mam.\*

37: em\_hgt\_vrt.\*

38: em\_sv.\*

39: em\_higo\_hum.\*

40: em\_higo\_mus.\*

41: em\_higo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	53	41.4	4229	9	AF234618	AF234618 Homo sapi
c 2	53	41.4	157284	9	AC072051	AC072051 Homo sapi
c 3	53	41.4	18092	2	AC027524	AC027524 Homo sapi
c 4	21	16.4	6182	6	AX346917	AX346917 Sequence
5	21	16.4	194622	2	EX296551	EX296551 Danio rer
6	20	15.6	47577	3	AF396436	AF396436 Tetrahyme
7	19	14.8	316	11	G30073	G30073 human STS S
8	19	14.8	482	6	AX337914	AX337914 Sequence
c 9	19	14.8	34372	3	U58727	U58727 Caenorhabdi
c 10	19	14.8	37490	9	AC026356	AC026356 Homo sapi
c 11	19	14.8	74350	8	AP002543	AP002543 Arabidops
c 12	19	14.8	91200	9	AP001152	AP001152 Homo sapi
c 13	19	14.8	107365	4	AC087861	AC087861 Felis cat
c 14	19	14.8	147556	2	AC011007	AC011007 Homo sapi
c 15	19	14.8	153875	9	AC003682	AC003682 Homo sapi
c 16	19	14.8	156394	2	AC026558	AC026558 Homo sapi
c 17	19	14.8	169529	9	AC064834	AC064834 Homo sapi
c 18	19	14.8	173235	2	AC023756	AC023756 Homo sapi
c 19	19	14.8	179269	2	AC116374	AC116374 Mus muscu
c 20	19	14.8	183395	2	AP004843	AP004843 Oryza sat
c 21	19	14.8	186291	9	AL954319	AL954319 Human DNA
c 22	19	14.8	188324	2	EX530094	EX530094 Homo sapi
c 23	19	14.8	201398	2	AC103354	AC103354 Mus muscu
c 24	19	14.8	220480	2	AC023973	AC023973 Homo sapi
c 25	19	14.8	245032	2	AC109749	AC109749 Rattus no
c 26	19	14.8	247582	2	AC115339	AC115339 Rattus no
c 27	19	14.8	257556	2	AC098051	AC098051 Rattus no
c 28	19	14.8	258646	2	AC096239	AC096239 Rattus no
c 29	19	14.8	306037	1	AB014017	AB014017 Buchnera
30	18	14.1	2000	6	AX655811	AX655811 Sequence
31	18	14.1	5232	10	AK122448	AK122448 Mus muscu
32	18	14.1	63807	2	AC087299	AC087299 Homo sapi
33	18	14.1	71256	2	AC105011	AC105011 Homo sapi
34	18	14.1	79526	9	AC025244	AC025244 Homo sapi
35	18	14.1	82872	9	AC093160	AC093160 Homo sapi
c 36	18	14.1	93614	9	AL139294	AL139294 Human DNA
c 37	18	14.1	110000	2	AC105643	Continuation (5 of
c 38	18	14.1	110096	10	AL928981	AL928981 Mouse DNA
c 39	18	14.1	120664	10	AL928981	AL928981 Mouse DNA
c 40	18	14.1	127125	9	AC107020	AC107020 Homo sapi
c 41	18	14.1	129241	9	AL353660	AL353660 Human DNA
c 42	18	14.1	140640	9	AC113165	AC113165 Homo sapi
c 43	18	14.1	143060	2	AC144350	AC144350 Homo sapi
c 44	18	14.1	143961	2	AP005862	AP005862 Oryza sat
c 45	18	14.1	155034	10	AL662803	AL662803 Mouse DNA
c 46	18	14.1	155332	9	AC109581	AC109581 Homo sapi
c 47	18	14.1	156801	5	AL929467	AL929467 Zebrfish
c 48	18	14.1	159056	8	CNS08CD7	AL954853 Oryza sat
c 49	18	14.1	160178	9	AL162584	AL162584 Human DNA
c 50	18	14.1	160654	2	AC011879	AC011879 Homo sapi
c 51	18	14.1	161797	2	AC027570	AC027570 Homo sapi
c 52	18	14.1	165036	9	AC133010	AC133010 Homo sapi
c 53	18	14.1	168108	9	AC104119	AC104119 Homo sapi
c 54	18	14.1	168473	9	AL512324	AL512324 Human DNA
c 55	18	14.1	169350	2	AC024332	AC024332 Homo sapi
c 56	18	14.1	169660	2	AC126028	AC126028 Mus muscu
c 57	18	14.1	169998	6	AR225961	AR225961 Sequence
c 58	18	14.1	171432	2	AC102322	AC102322 Mus muscu
c 59	18	14.1	171432	2	AC102322	AC102322 Mus muscu
c 60	18	14.1	172097	2	AC143354	AC143354 Homo sapi
c 61	18	14.1	174099	2	AC143333	AC143333 Homo sapi
c 62	18	14.1	177998	9	AC006977	AC006977 Homo sapi
c 63	18	14.1	185470	9	AC010947	AC010947 Homo sapi
c 64	18	14.1	185535	9	AC022272	AC022272 Homo sapi
c 65	18	14.1	188820	2	AC144070	AC144070 Macaca mu
	18	14.1	193467	2	AC108423	AC108423 Mus muscu

66	18	14.1	195636	2	AC139391	AC139391 Rattus no	139	17	13.3	44079	9	AC138626	AC138626 Homo sapi
c 67	18	14.1	197496	6	AX644070	Sequence	c 140	17	13.3	52754	2	AC099928	AC099928 Mus muscu
c 68	18	14.1	197496	9	AF288738	Homo sapi	c 141	17	13.3	53963	6	AX695428	Sequence
c 69	18	14.1	198344	2	AC087872	Mus muscu	c 142	17	13.3	61812	9	AL591395	Human DNA
c 70	18	14.1	199301	2	EX088589	Danio rer	c 143	17	13.3	65158	2	AC100429	Mus muscu
c 71	18	14.1	199632	2	AC099703	AC099703 Mus muscu	c 144	17	13.3	66665	2	AC116006	AC116006 Homo sapi
c 72	18	14.1	202074	2	AC101884	AC101884 Mus muscu	c 145	17	13.3	69775	9	AC006347	AC006347 Homo sapi
c 73	18	14.1	203010	9	AC006530	AC006530 Homo sapi	c 146	17	13.3	86380	10	MMU131017	MMU131017 Mus muscu
c 74	18	14.1	203980	2	AC134058	AC134058 Rattus no	c 147	17	13.3	86574	9	H8833B7	AC008637 Human DNA
c 75	18	14.1	205515	2	AC118905	AC118905 Rattus no	c 148	17	13.3	88087	5	AL107078	AC110708 Rattus no
c 76	18	14.1	213996	9	AC019115	AC019115 Homo sapi	c 149	17	13.3	92440	2	AL954729	AL954729 Zebrafish
c 77	18	14.1	215816	2	AC134937	AC134937 Rattus no	c 150	17	13.3	93398	2	AP001863	AP001863 Homo sapi
c 78	18	14.1	215860	2	AC125959	AC125959 Rattus no	c 151	17	13.3	95129	9	AC003099	AC003099 Homo sapi
c 79	18	14.1	218874	10	AC105600	AC105600 Rattus no	c 152	17	13.3	96595	6	AX695827	AX695827 Sequence
c 80	18	14.1	223363	2	AC120700	AC120700 Rattus no	c 153	17	13.3	98239	2	AC131553	Continuation (4 of
c 81	18	14.1	224210	2	AC113109	AC113109 Mus muscu	c 154	17	13.3	103139	9	AL139413	Human DNA
c 82	18	14.1	226937	2	AC117052	AC117052 Rattus no	c 155	17	13.3	104757	2	AP000625	AP000625 Homo sapi
c 83	18	14.1	230936	2	AC116708	AC116708 Mus muscu	c 156	17	13.3	108388	9	AP001821	AP001821 Homo sapi
c 84	18	14.1	233247	2	AC134936	AC134936 Rattus no	c 157	17	13.3	108587	2	EX248309	EX248309 Homo sapi
c 85	18	14.1	233307	2	AC097932	AC097932 Rattus no	c 158	17	13.3	110000	2	AC095225	Continuation (3 of
c 86	18	14.1	236425	2	AC127803	AC127803 Rattus no	c 159	17	13.3	110000	2	AC129176	Continuation (3 of
c 87	18	14.1	238123	2	AC136571	AC136571 Rattus no	c 160	17	13.3	110000	2	AL359456	AL359456 Homo sapi
c 88	18	14.1	238226	2	AC111396	AC111396 Rattus no	c 161	17	13.3	110000	2	EX004982	EX004982 Homo sapi
c 89	18	14.1	240918	2	AC137254	AC137254 Rattus no	c 162	17	13.3	112875	9	AP000638	AP000638 Homo sapi
c 90	18	14.1	243968	2	AC130131	AC130131 Rattus no	c 163	17	13.3	113168	9	H87779E11	AL121675 Human DNA
c 91	18	14.1	245824	2	AC135970	AC135970 Rattus no	c 164	17	13.3	114149	9	AC022083	AC022083 Homo sapi
c 92	18	14.1	247462	2	AC134478	AC134478 Rattus no	c 165	17	13.3	116260	9	AC025573	AC025573 Homo sapi
c 93	18	14.1	248014	2	AC112043	AC112043 Rattus no	c 166	17	13.3	118011	2	AL808138	AL808138 Mus muscu
c 94	18	14.1	248284	2	AC106584	AC106584 Rattus no	c 167	17	13.3	122903	9	AC005833	AC005833 Homo sapi
c 95	18	14.1	252721	2	AC095946	AC095946 Rattus no	c 168	17	13.3	123993	9	AC018802	AC018802 Homo sapi
c 96	18	14.1	252876	2	AC091272	AC091272 Mus muscu	c 169	17	13.3	126495	9	AL450310	AL450310 Human DNA
c 97	18	14.1	253287	2	AC137417	AC137417 Rattus no	c 170	17	13.3	126781	10	AL732622	AL732622 Mouse DNA
c 98	18	14.1	272652	2	AC109709	AC109709 Rattus no	c 171	17	13.3	128910	10	AC125308	AC125308 Mus muscu
c 99	18	14.1	273797	2	AC115397	AC115397 Rattus no	c 172	17	13.3	129738	9	AC074239	AC074239 Homo sapi
c 100	18	14.1	277164	2	AC110366	AC110366 Rattus no	c 173	17	13.3	130607	9	AC004875	AC004875 Homo sapi
c 101	18	14.1	280334	7	AF399011	AF399011 Pseudomon	c 174	17	13.3	131971	10	AL845434	AL845434 Mouse DNA
c 102	18	14.1	282132	2	AC109976	AC109976 Rattus no	c 175	17	13.3	135455	9	AL513666	AL513666 Human DNA
c 103	18	14.1	323991	2	AC098512	AC098512 Rattus no	c 176	17	13.3	136332	2	AC131041	AC131041 Mus muscu
c 104	18	14.1	343858	2	AC095580	AC095580 Rattus no	c 177	17	13.3	136725	9	AC026369	AC026369 Homo sapi
c 105	17	13.3	216	9	HSP40PHOX5	US0724 Human P40-P	c 178	17	13.3	136981	2	AC132586	AC132586 Mus muscu
c 106	17	13.3	554	11	G88440	G88440 S08F068RF	c 179	17	13.3	137478	5	AL929458	AL929458 Zebrafish
c 107	17	13.3	687	8	AF180365	AF180365 Hieracium	c 180	17	13.3	140325	2	AC124620	AC124620 Mus muscu
c 108	17	13.3	775	10	AY169688	AY169688 Mus muscu	c 181	17	13.3	142678	2	AC131890	AC131890 Ateleterix
c 109	17	13.3	790	10	MDIYGDP	Z17336 M.domesticu	c 182	17	13.3	142895	9	AC090615	AC090615 Homo sapi
c 110	17	13.3	875	6	AF180364	AF180364 Hieracium	c 183	17	13.3	143841	9	AC008964	AC008964 Homo sapi
c 111	17	13.3	936	6	AX142537	AX142537 Sequence	c 184	17	13.3	144290	2	AC034251	AC034251 Homo sapi
c 112	17	13.3	1020	6	AX658907	AX658907 Sequence	c 185	17	13.3	144419	2	AC079968	AC079968 Homo sapi
c 113	17	13.3	1047	9	BT007346	BT007346 Homo sapi	c 186	17	13.3	145086	2	AC021976	AC021976 Homo sapi
c 114	17	13.3	1056	9	AB025220	AB025220 Homo sapi	c 187	17	13.3	147405	2	AC118576	AC118576 Didelphis
c 115	17	13.3	1103	6	AX151683	AX151683 Sequence	c 188	17	13.3	147896	10	AL607145	AL607145 Mouse DNA
c 116	17	13.3	1245	9	HSP40PHOX	X77094 H.sapiens m	c 189	17	13.3	148282	9	AC024649	AC024649 Homo sapi
c 117	17	13.3	1261	6	BD092311	BD092311 Identific	c 190	17	13.3	148468	2	AC101715	AC101715 Mus muscu
c 118	17	13.3	1301	9	AB025219	AB025219 Homo sapi	c 191	17	13.3	148510	10	AC113889	AC113889 Rattus no
c 119	17	13.3	1665	9	H53785	X96895 H.sapiens s	c 192	17	13.3	148592	2	AC125190	AC125190 Mus muscu
c 120	17	13.3	1713	9	BC002798	BC002798 Homo sapi	c 193	17	13.3	148900	2	AC025404	AC025404 Homo sapi
c 121	17	13.3	1807	6	AX399856	AX399856 Sequence	c 194	17	13.3	150350	9	CNS01DW	AL138539 Human chr
c 122	17	13.3	1846	6	AX399855	AX399855 Sequence	c 195	17	13.3	150962	2	AC116375	AC116375 Mus muscu
c 123	17	13.3	2692	6	AX695906	AX695906 Sequence	c 196	17	13.3	151368	9	AC040171	AC040171 Homo sapi
c 124	17	13.3	3356	1	AF269880	AF269880 Staphyloc	c 197	17	13.3	151580	2	AP005441	AP005441 Oryza sat
c 125	17	13.3	3356	6	AX145198	AX145198 Sequence	c 198	17	13.3	151580	2	AP005441	AP005441 Oryza sat
c 126	17	13.3	3376	1	AF269403	AF269403 Staphyloc	c 199	17	13.3	151884	9	AC108103	AC108103 Homo sapi
c 127	17	13.3	3376	6	AX144723	AX144723 Sequence	c 200	17	13.3	152560	9	AC011339	AC011339 Homo sapi
c 128	17	13.3	10351	1	AB000560	AB000560 Helicobac	c 201	17	13.3	152780	2	AC139961	AC139961 Rattus no
c 129	17	13.3	16918	6	AX346519	AX346519 Sequence	c 202	17	13.3	153223	2	AC018548	AC018548 Homo sapi
c 130	17	13.3	18347	3	CRT22H2	Z81595 Caenorhabdi	c 203	17	13.3	153775	2	AC011943	AC011943 Homo sapi
c 131	17	13.3	19489	2	AC015314	AC015314 Drosophil	c 204	17	13.3	157053	2	AC131642	AC131642 Rattus no
c 132	17	13.3	33822	2	AC017206	AC017206 Drosophil	c 205	17	13.3	158317	2	AL390724	AL390724 Homo sapi
c 133	17	13.3	36852	3	L14429	L14429 Caenorhabdi	c 206	17	13.3	158525	2	AC122208	AC122208 Mus muscu
c 134	17	13.3	36875	9	AC079825	AC079825 Homo sapi	c 207	17	13.3	159240	9	AC104995	AC104995 Homo sapi
c 135	17	13.3	39212	9	H5N21F1	Z94162 Human DNA s	c 208	17	13.3	159486	10	AC123814	AC123814 Mus muscu
c 136	17	13.3	40304	6	AX695905	AX695905 Sequence	c 209	17	13.3	160956	2	AC013637	AC013637 Homo sapi
c 137	17	13.3	41482	3	AF106577	AF106577 Caenorhab	c 210	17	13.3	161633	2	AL592388	AL592388 Homo sapi
c 138	17	13.3	41645	3	L11247	L11247 Caenorhabdi	c 211	17	13.3	161815	9	AC092015	AC092015 Homo sapi



212	17	13.3	161984	10	AC104296	AC104296 Mus muscu	c 285	17	13.3	191155	2	AC117803	AC117803 Mus muscu
c 213	17	13.3	162215	10	AL731773	AL731773 Mouse DNA	c 286	17	13.3	191405	2	AC139454	AC139454 Homo sapi
c 214	17	13.3	162533	2	EX537304	EX537304 Danio rer	c 287	17	13.3	191896	9	AC009481	AC009481 Homo sapi
c 215	17	13.3	162876	3	AC008212	AC008212 Drosophi	c 288	17	13.3	191911	9	AC009865	AC009865 Homo sapi
c 216	17	13.3	163009	9	AC011884	AC011884 Homo sapi	c 289	17	13.3	193363	2	AC090432	AC090432 Mus muscu
c 217	17	13.3	163360	2	AC021598	AC021598 Homo sapi	c 290	17	13.3	193495	2	AC122100	AC122100 Rattus no
c 218	17	13.3	163588	2	AC142022	AC142022 Rattus no	c 291	17	13.3	193592	2	AC124095	AC124095 Mus muscu
c 219	17	13.3	163672	2	AC084256	AC084256 Homo sapi	c 292	17	13.3	194158	9	AC021127	AC021127 Homo sapi
c 220	17	13.3	164815	2	AC013668	AC013668 Homo sapi	c 293	17	13.3	194290	2	AC137228	AC137228 Rattus no
c 221	17	13.3	165725	2	AC102041	AC102041 Mus muscu	c 294	17	13.3	194317	2	AC137232	AC137232 Rattus no
c 222	17	13.3	165764	9	AC116917	AC116917 Homo sapi	c 295	17	13.3	195282	2	AC022453	AC022453 Mus muscu
c 223	17	13.3	165890	10	AC121891	AC121891 Mus muscu	c 296	17	13.3	195931	9	AC010633	AC010633 Homo sapi
c 224	17	13.3	166007	9	CNS01DRC	AL117186 Human chr	c 297	17	13.3	197326	2	AC122380	AC122380 Mus muscu
c 225	17	13.3	166543	3	AC005813	AC005813 Drosophi	c 298	17	13.3	197660	10	AL589650	AL589650 Mouse DNA
c 226	17	13.3	166742	2	AC109221	AC109221 Mus muscu	c 299	17	13.3	197978	10	AC114821	AC114821 Mus muscu
c 227	17	13.3	167559	9	AC090042	AC090042 Homo sapi	c 300	17	13.3	198031	10	AL645470	AL645470 Mouse DNA
c 228	17	13.3	168357	9	AC096540	AC096540 Homo sapi	c 301	17	13.3	198309	2	AC120518	AC120518 Sus scrof
c 229	17	13.3	168965	9	AC012101	AC012101 Homo sapi	c 302	17	13.3	198966	2	AC120873	AC120873 Mus muscu
c 230	17	13.3	169048	9	AL158011	AL158011 Human DNA	c 303	17	13.3	199757	2	AC112985	AC112985 Mus muscu
c 231	17	13.3	169420	2	AC118641	AC118641 Mus muscu	c 304	17	13.3	200341	2	AC130577	AC130577 Rattus no
c 232	17	13.3	169468	9	AC006928	AC006928 Homo sapi	c 305	17	13.3	200341	2	AC138403	AC138403 Mus muscu
c 233	17	13.3	169942	2	AC124793	AC124793 Homo sapi	c 306	17	13.3	200760	5	AL528872	AL528872 Zebrafish
c 234	17	13.3	170206	2	AC138258	AC138258 Homo sapi	c 307	17	13.3	201132	2	EX088555	EX088555 Mus muscu
c 235	17	13.3	170740	9	HSBA90K6	AL121902 Human DNA	c 308	17	13.3	201142	10	AL670035	AL670035 Mouse DNA
c 236	17	13.3	170763	9	AC009196	AC009196 Homo sapi	c 309	17	13.3	201495	10	AL591884	AL591884 Mouse DNA
c 237	17	13.3	170829	2	AP001316	AP001316 Homo sapi	c 310	17	13.3	204498	5	AL831791	AL831791 Zebrafish
c 238	17	13.3	171349	2	AC020735	AC020735 Homo sapi	c 311	17	13.3	204662	2	AC127364	AC127364 Mus muscu
c 239	17	13.3	171316	2	AC104909	AC104909 Mus muscu	c 312	17	13.3	205202	2	AC139347	AC139347 Mus muscu
c 240	17	13.3	172388	2	AC137258	AC137258 Rattus no	c 313	17	13.3	205560	2	AC144630	AC144630 Mus muscu
c 241	17	13.3	173043	2	AC124792	AC124792 Homo sapi	c 314	17	13.3	205694	2	AC098183	AC098183 Rattus no
c 242	17	13.3	173933	9	AP002022	AP002022 Homo sapi	c 315	17	13.3	207183	9	AP001533	AP001533 Homo sapi
c 243	17	13.3	174009	9	AC007842	AC007842 Homo sapi	c 316	17	13.3	207199	2	AC036183	AC036183 Homo sapi
c 244	17	13.3	174287	9	AC100840	AC100840 Homo sapi	c 317	17	13.3	207925	2	EX537115	EX537115 Danio rer
c 245	17	13.3	174409	2	AL359754	AL359754 Homo sapi	c 318	17	13.3	208166	2	AC131980	AC131980 Mus muscu
c 246	17	13.3	175258	2	AC017098	AC017098 Homo sapi	c 319	17	13.3	208651	10	AC090493	AC090493 Genomic S
c 247	17	13.3	176329	2	AC119047	AC119047 Rattus no	c 320	17	13.3	208806	2	AC099941	AC099941 Mus muscu
c 248	17	13.3	176699	2	AC122967	AC122967 Rattus no	c 321	17	13.3	210988	2	AC139449	AC139449 Rattus no
c 249	17	13.3	178341	2	AC067846	AC067846 Homo sapi	c 322	17	13.3	215159	2	AC133516	AC133516 Mus muscu
c 250	17	13.3	178611	2	AC024631	AC024631 Homo sapi	c 323	17	13.3	215781	2	AC098331	AC098331 Rattus no
c 251	17	13.3	179141	2	AC019091	AC019091 Homo sapi	c 324	17	13.3	216133	2	AC087153	AC087153 Mus muscu
c 252	17	13.3	179383	2	AC141997	AC141997 Rattus no	c 325	17	13.3	219635	2	AL390207	AL390207 Homo sapi
c 253	17	13.3	179696	2	AC135115	AC135115 Mus muscu	c 326	17	13.3	220599	2	AL357792	AL357792 Homo sapi
c 254	17	13.3	179816	2	AC102674	AC102674 Mus muscu	c 327	17	13.3	222922	2	AC106497	AC106497 Rattus no
c 255	17	13.3	179943	10	AC140424	AC140424 Mus muscu	c 328	17	13.3	222932	2	AC137869	AC137869 Mus muscu
c 256	17	13.3	179992	2	AC091251	AC091251 Mus muscu	c 329	17	13.3	223897	2	AC117154	AC117154 Rattus no
c 257	17	13.3	180280	2	AC139054	AC139054 Mus muscu	c 330	17	13.3	224128	2	AC125046	AC125046 Mus muscu
c 258	17	13.3	180286	9	AC010543	AC010543 Homo sapi	c 331	17	13.3	225645	2	AC117662	AC117662 Mus muscu
c 259	17	13.3	180418	9	AC093875	AC093875 Homo sapi	c 332	17	13.3	225846	2	AC109040	AC109040 Rattus no
c 260	17	13.3	180754	9	AC069366	AC069366 Homo sapi	c 333	17	13.3	225978	2	AC140462	AC140462 Mus muscu
c 261	17	13.3	180905	2	AC011688	AC011688 Homo sapi	c 334	17	13.3	226681	2	AC106383	AC106383 Rattus no
c 262	17	13.3	181103	10	AL596084	AL596084 Mouse DNA	c 335	17	13.3	227372	2	AC119847	AC119847 Mus muscu
c 263	17	13.3	181238	2	AC116937	AC116937 Bos tauru	c 336	17	13.3	227453	2	AC106681	AC106681 Rattus no
c 264	17	13.3	182080	3	AC012373	AC012373 Drosophi	c 337	17	13.3	227576	2	AC134040	AC134040 Rattus no
c 265	17	13.3	182176	2	AC109501	AC109501 Mus muscu	c 338	17	13.3	228075	2	AC134523	AC134523 Mus muscu
c 266	17	13.3	183370	3	AC008213	AC008213 Drosophi	c 339	17	13.3	228276	2	AC120908	AC120908 Rattus no
c 267	17	13.3	183378	2	AL607151	AL607151 Mus muscu	c 340	17	13.3	228696	2	AC112425	AC112425 Rattus no
c 268	17	13.3	183446	10	AC108946	AC108946 Mus muscu	c 341	17	13.3	229126	2	AC105803	AC105803 Rattus no
c 269	17	13.3	183719	2	AC068842	AC068842 Homo sapi	c 342	17	13.3	229363	2	AC094940	AC094940 Rattus no
c 270	17	13.3	185098	2	AC140265	AC140265 Mus muscu	c 343	17	13.3	230417	2	AC120447	AC120447 Rattus no
c 271	17	13.3	185257	2	AC117777	AC117777 Mus muscu	c 344	17	13.3	230956	2	AC105782	AC105782 Rattus no
c 272	17	13.3	185349	2	AC105173	AC105173 Mus muscu	c 345	17	13.3	231198	2	AC126892	AC126892 Rattus no
c 273	17	13.3	185784	9	AC096638	AC096638 Homo sapi	c 346	17	13.3	232343	2	AC115710	AC115710 Mus muscu
c 274	17	13.3	186621	10	AL773517	AL773517 Mouse DNA	c 347	17	13.3	233231	2	AC099137	AC099137 Rattus no
c 275	17	13.3	186927	9	AC078844	AC078844 Homo sapi	c 348	17	13.3	233454	9	AC005703	AC005703 Homo sapi
c 276	17	13.3	187321	2	AC092510	AC092510 Papio anu	c 349	17	13.3	234411	10	MMAJ2971	MMAJ2971 Mus muscu
c 277	17	13.3	188579	2	AC107316	AC107316 Mus muscu	c 350	17	13.3	236464	2	AC131833	AC131833 Rattus no
c 278	17	13.3	188818	2	AC098674	AC098674 Homo sapi	c 351	17	13.3	236472	2	AC118297	AC118297 Rattus no
c 279	17	13.3	188926	9	AC012182	AC012182 Homo sapi	c 352	17	13.3	237359	2	AC126505	AC126505 Rattus no
c 280	17	13.3	189167	9	AC116003	AC116003 Homo sapi	c 353	17	13.3	237450	2	AC095092	AC095092 Rattus no
c 281	17	13.3	189715	2	AC025558	AC025558 Homo sapi	c 354	17	13.3	238404	2	AC139759	AC139759 Mus muscu
c 282	17	13.3	190123	10	AC122876	AC122876 Mus muscu	c 355	17	13.3	238487	2	AC136379	AC136379 Rattus no
c 283	17	13.3	190520	2	AC138271	AC138271 Homo sapi	c 356	17	13.3	238556	2	EX470214	EX470214 Danio rer
c 284	17	13.3	190787	2	AC105171	AC105171 Mus muscu	c 357	17	13.3	238844	2	AC108261	AC108261 Rattus no

C 358	17	13.3	241404	2	AC099365	AC099365 Rattus no	431	15	12.5	1485	8	SSP429730	AJ429730 Scaevola
C 359	17	13.3	243220	2	AC135527	AC135527 Rattus no	432	16	12.5	1563	6	AX488920	AX488920 Sequence
C 360	17	13.3	245485	2	AC135745	AC135745 Rattus no	C 433	16	12.5	1569	5	EC043857	EC043857 Xenopus 1
C 361	17	13.3	245937	2	AC133984	AC133984 Rattus no	434	16	12.5	1685	6	EC00016	EC00016 DNA coding
C 362	17	13.3	248899	2	AC095985	AC095985 Rattus no	435	16	12.5	1685	6	EC01203	EC01203 DNA sequence
C 363	17	13.3	249223	2	AC134066	AC134066 Rattus no	436	16	12.5	1709	6	AI2109	AI2109 oligonucleo
C 364	17	13.3	249834	2	AC132449	AC132449 Mus muscu	437	16	12.5	1765	3	AY070917	AY070917 Drosophil
C 365	17	13.3	250415	2	AC097418	AC097418 Rattus no	C 438	16	12.5	1872	9	AK027181	AK027181 Homo sapi
C 366	17	13.3	252313	2	AC127110	AC127110 Rattus no	C 439	16	12.5	1923	6	AR165242	AR165242 Sequence
C 367	17	13.3	254336	2	AC096105	AC096105 Rattus no	C 440	16	12.5	1938	6	AX481462	AX481462 Sequence
C 368	17	13.3	255156	2	AC111130	AC111130 Mus muscu	C 441	16	12.5	1938	6	EC39423	EC39423 Novel membr
C 369	17	13.3	255706	2	AC094186	AC094186 Rattus no	442	16	12.5	2022	9	HSIFNA4B	HSIFNA4B
C 370	17	13.3	255881	2	AC102994	AC102994 Rattus no	C 443	16	12.5	2076	2	AX429950	AX429950 Human inter
C 371	17	13.3	255924	2	AC096518	AC096518 Rattus no	C 444	16	12.5	2118	6	EC39430	EC39430 Novel membr
C 372	17	13.3	257391	2	AC097400	AC097400 Rattus no	C 445	16	12.5	2118	9	AF131284	AF131284 Homo sapi
C 373	17	13.3	258332	3	AE003756	AE003756 Drosophil	C 446	16	12.5	2170	10	BC037030	BC037030 Mus muscu
C 374	17	13.3	258545	2	AC135666	AC135666 Rattus no	C 447	16	12.5	2181	3	AF136828	AF136828 Schlitoso
C 375	17	13.3	258830	2	AC095398	AC095398 Rattus no	C 448	16	12.5	2206	5	GGU60297	GGU60297 Gallus gall
C 376	17	13.3	262015	2	AC095852	AC095852 Rattus no	449	16	12.5	2299	8	GFU417493	GFU417493 Gibberell
C 377	17	13.3	262843	2	AC097805	AC097805 Rattus no	450	16	12.5	2424	1	ECOMSDNA	ECOMSDNA
C 378	17	13.3	266132	2	AC099138	AC099138 Rattus no	451	16	12.5	2501	5	AF141607	AF141607 Salmo sal
C 379	17	13.3	266587	2	AC114098	AC114098 Rattus no	C 452	16	12.5	2612	9	AF141606	AF141606 Salmo sal
C 380	17	13.3	266926	2	AC112028	AC112028 Rattus no	C 453	16	12.5	2623	5	AB021227	AB021227 Homo sapi
C 381	17	13.3	267506	2	AC097771	AC097771 Rattus no	C 454	16	12.5	2868	10	AF277093	AF277093 Mus muscu
C 382	17	13.3	268694	2	AC129752	AC129752 Rattus no	C 455	16	12.5	3017	8	YSCRP5	YSCRP5
C 383	17	13.3	270667	2	AC095697	AC095697 Rattus no	456	16	12.5	3168	9	AK090596	AK090596 Homo sapi
C 384	17	13.3	273219	2	EX005325	EX005325 Danio rer	C 457	16	12.5	3315	9	EC021253	EC021253 Homo sapi
C 385	17	13.3	275924	2	AC116233	AC116233 Rattus no	C 458	16	12.5	3343	9	AK091912	AK091912 Homo sapi
C 386	17	13.3	278514	2	AC095674	AC095674 Rattus no	C 459	16	12.5	3375	6	BD159667	BD159667 Primer to
C 387	17	13.3	280138	2	AC095813	AC095813 Rattus no	C 460	16	12.5	3375	9	AK022845	AK022845 Homo sapi
C 388	17	13.3	282610	2	AC096244	AC096244 Rattus no	C 461	16	12.5	3381	9	HSM801540	HSM801540 Homo sapi
C 389	17	13.3	286363	2	AC095687	AC095687 Rattus no	C 462	16	12.5	3381	9	HSB805198	HSB805198 Homo sapi
C 390	17	13.3	292977	14	AF369029	AF369029 White spo	C 463	16	12.5	3416	8	SCYBR237W	SCYBR237W
C 391	17	13.3	293576	2	AC097830	AC097830 Rattus no	C 464	16	12.5	3449	6	AX135358	AX135358 Sequence
C 392	17	13.3	295052	2	AC111478	AC111478 Rattus no	C 465	16	12.5	3470	6	AX430949	AX430949 Sequence
C 393	17	13.3	300242	1	AE016745	AE016745 Staphyloc	C 466	16	12.5	3490	10	AF097544	AF097544 Mus muscu
C 394	17	13.3	300988	2	AC096249	AC096249 Rattus no	467	16	12.5	3516	3	PBB420907	PBB420907 Plasmid
C 395	17	13.3	301888	2	AC094881	AC094881 Rattus no	C 468	16	12.5	3748	3	AF005034	AF005034 Spodopter
C 396	17	13.3	305107	6	AX151396	AX151396 Sequence	C 469	16	12.5	3748	6	AR235899	AR235899 Sequence
C 397	17	13.3	305107	14	AF332093	AF332093 White spo	C 470	16	12.5	3751	5	CHKGH6A	CHKGH6A
C 398	17	13.3	305109	3	AE003509	AE003509 Drosophil	471	16	12.5	3852	4	D9006585	D9006585 Bos taurus
C 399	17	13.3	307287	14	AF440570	AF440570 Shrimp wh	472	16	12.5	4225	9	EC00XRETB	EC00XRETB
C 400	17	13.3	308232	2	AC1392004	AC1392004 Homo sapi	C 473	16	12.5	4263	9	D84103	D84103 Homo sapien
C 401	17	13.3	310197	2	AC114599	AC114599 Mus muscu	C 474	16	12.5	4312	9	BC047614	BC047614 Homo sapi
C 402	17	13.3	313650	2	AC131995	AC131995 Mus muscu	C 475	16	12.5	4320	9	HSU60325	HSU60325 Human DNA p
C 403	17	13.3	327738	2	AC104922	AC104922 Mus muscu	C 476	16	12.5	4440	6	AX034450	AX034450 Sequence
C 404	17	13.3	328868	2	AC118955	AC118955 Rattus no	C 477	16	12.5	4465	9	MIHSDNAPL	MIHSDNAPL
C 405	16	12.5	47	6	AR291937	AR291937 Sequence	C 478	16	12.5	4503	9	BC050559	BC050559 Homo sapi
C 406	16	12.5	51	10	MMU90480	MMU90480 Mus musculu	C 479	16	12.5	4563	1	SAU292927	SAU292927 Staphyloc
C 407	16	12.5	222	11	G24200	G24200 human STS W	C 480	16	12.5	4563	1	SAU292927	SAU292927 Staphyloc
C 408	16	12.5	344	5	AB100555	AB100555 Tylochrom	C 481	16	12.5	6160	6	AX344268	AX344268 Sequence
C 409	16	12.5	380	6	AX070390	AX070390 Sequence	C 482	16	12.5	6160	6	AX348665	AX348665 Sequence
C 410	16	12.5	385	6	AX332693	AX332693 Sequence	C 483	16	12.5	7186	1	AF102623	AF102623 Methanosa
C 411	16	12.5	385	6	AX336214	AX336214 Sequence	C 484	16	12.5	7920	1	AE010010	AE010010 Streptoco
C 412	16	12.5	599	11	BV065135	BV065135 S212P6036	C 485	16	12.5	8301	6	BD016747	BD016747 Protein P
C 413	16	12.5	598	11	BV041105	BV041105 S212P6027	C 486	16	12.5	8301	10	AF169411	AF169411 Rattus no
C 414	16	12.5	666	11	G56448	G56448 SHGC-101991	487	16	12.5	8630	5	AF141605	AF141605 Salmo sal
C 415	16	12.5	735	6	I08307	I08307 Sequence 1	488	16	12.5	9293	5	AF278717	AF278717 Salmo sal
C 416	16	12.5	753	6	AX592877	AX592877 Sequence	489	16	12.5	9949	3	CELINABG	CELINABG
C 417	16	12.5	756	6	AX592893	AX592893 Sequence	C 490	16	12.5	10662	2	AC020156	AC020156 Drosophil
C 418	16	12.5	758	6	AX392887	AX392887 Sequence	491	16	12.5	11334	1	EC067RTDM	EC067RTDM
C 419	16	12.5	763	6	I08302	I08302 Sequence 3	C 492	16	12.5	11461	1	AE008587	AE008587 Rickettsi
C 420	16	12.5	763	6	I08308	I08308 Sequence 3	493	16	12.5	11730	3	CEUI0413	CEUI0413 Caenorhabdi
C 421	16	12.5	857	10	AF425090	AF425090 Mus muscu	C 494	16	12.5	11959	8	AF100661	AF100661 Caenorhab
C 422	16	12.5	876	6	I08301	I08301 Sequence 1	C 495	16	12.5	12679	3	AF210617	AF210617 Zea mays
C 423	16	12.5	896	6	AX592889	AX592889 Sequence	496	16	12.5	15993	2	AC138332	AC138332
C 424	16	12.5	958	6	BD078471	BD078471 101 human	C 497	16	12.5	17935	10	AB011006	AB011006 Rattus no
C 425	16	12.5	977	6	AO2076	AO2076 H.sapiens m	C 498	16	12.5	18978	2	AC114431	AC114431
C 426	16	12.5	977	9	HUMFNAM1	HUMFNAM1	C 499	16	12.5	19514	2	AC018009	AC018009 Drosophil
C 427	16	12.5	1001	11	CNS06F90	CNS06F90	C 500	16	12.5	21304	9	AF497906	AF497906 Homo sapi
C 428	16	12.5	1092	10	AF425086	AF425086 Mus muscu	C 501	16	12.5	25224	1	AF513500	AF513500 Mycobacte
C 429	16	12.5	1122	3	AB003722	AB003722 Sternaspi	C 502	16	12.5	25823	3	AC006632	AC006632 Caenorhab
C 430	16	12.5	1130	10	MMNF2MR	MMNF2MR	C 503	16	12.5	29214	3	U80445	U80445 Caenorhabdi

504	16	12.5	29688	3	CRC09G5	Z46791 Caenorhabdi	c 577	16	12.5	100000	9	AP000100 Homo sapi
505	16	12.5	31310	3	U80837	U80837 Caenorhabdi	c 578	16	12.5	100000	9	AP000176 Homo sapi
506	16	12.5	33800	3	CEKX662	Z79604 Caenorhabdi	c 579	16	12.5	100300	2	AC114370 Homo sapi
507	16	12.5	34520	3	CEKX03D7	Z81562 Caenorhabdi	c 580	16	12.5	100301	9	HS137E23
508	16	12.5	34759	3	CSC34C12	Z46996 Caenorhabdi	c 581	16	12.5	100355	2	AL645692 Mus muscu
509	16	12.5	37447	3	CEP11A5	Z92830 Caenorhabdi	c 582	16	12.5	101270	9	HS483K16
510	16	12.5	38764	6	AX695911	AX695911 Sequence	c 583	16	12.5	101623	9	AL450389
511	16	12.5	38969	3	U51995	U51995 Caenorhabdi	c 584	16	12.5	102990	2	AP000801
512	16	12.5	39643	3	U41028	U41028 Caenorhabdi	c 585	16	12.5	103567	9	HS124C6
513	16	12.5	40129	9	AF006501	AF006501 Homo sapi	c 586	16	12.5	106376	9	AL138714 Human DNA
514	16	12.5	40302	3	CRC11G6	Z70204 Caenorhabdi	c 587	16	12.5	106516	1	ST1PPHCM2
515	16	12.5	42979	3	AF043701	AF043701 Caenorhab	c 588	16	12.5	107427	2	HS481A17
516	16	12.5	43106	7	AY050245	AY050245 Streptoco	c 589	16	12.5	108145	9	AP004649
517	16	12.5	45355	3	U41279	U41279 Caenorhabdi	c 590	16	12.5	109859	9	HS747E2
518	16	12.5	46101	2	AC140692	AC140692 Rattus no	c 591	16	12.5	110000	2	AC094428 1
519	16	12.5	46150	2	AC102146	AC102146 Clona sav	c 592	16	12.5	110000	2	AC094997 0
520	16	12.5	48666	9	AC106893	AC106893 Homo sapi	c 593	16	12.5	110000	2	AC095260 2
521	16	12.5	49580	2	AC136347	AC136347 Homo sapi	c 594	16	12.5	110000	2	AC095863 02
522	16	12.5	49737	2	AC103678	AC103678 Homo sapi	c 595	16	12.5	110000	2	AC097441 1
523	16	12.5	49918	2	AC142208	AC142208 Homo sapi	c 596	16	12.5	110000	2	AC106281 0
524	16	12.5	50097	1	AB014146	AB014146 Streptoco	c 597	16	12.5	110000	2	AC106281 1
525	16	12.5	51680	9	AF000261	AF000261 Homo sapi	c 598	16	12.5	110000	2	AC106281 2
526	16	12.5	51908	1	AB014160	AB014160 Streptoco	c 599	16	12.5	110000	2	AC107149 0
527	16	12.5	54488	8	AF359360	AF359360 Fusarium	c 600	16	12.5	110000	2	AC108249 3
528	16	12.5	56248	2	AC096086 3	Continuation (4 of	c 601	16	12.5	110000	2	AC109710 2
529	16	12.5	56307	9	AL512606	AL512606 Human DNA	c 602	16	12.5	110000	2	AC110829 4
530	16	12.5	57633	9	AC093382	AC093382 Homo sapi	c 603	16	12.5	110000	2	AC110930 3
531	16	12.5	59460	2	AC130298	AC130298 Homo sapi	c 604	16	12.5	110000	2	AC112872 3
532	16	12.5	59460	2	AC130298	AC130298 Homo sapi	c 605	16	12.5	110000	2	AC112886 0
533	16	12.5	62964	2	AC104959	AC104959 Homo sapi	c 606	16	12.5	110000	2	AC113693 3
534	16	12.5	62964	2	AC104959	AC104959 Homo sapi	c 607	16	12.5	110000	2	AC113693 4
535	16	12.5	63518	2	AC095260 4	Continuation (5 of	c 608	16	12.5	110000	2	AC114216 1
536	16	12.5	64216	2	AC101352	AC101352 Mus muscu	c 609	16	12.5	110000	2	AC114431 3
537	16	12.5	66822	9	AL606529	AL606529 Human DNA	c 610	16	12.5	110000	2	AC114431 6
538	16	12.5	67399	2	AC021388	AC021388 Homo sapi	c 611	16	12.5	110000	2	AC117319 0
539	16	12.5	67701	2	AC133629	AC133629 Homo sapi	c 612	16	12.5	110000	2	AC118411 2
540	16	12.5	68723	2	AC125941 3	Continuation (4 of	c 613	16	12.5	110000	2	AC119019 0
541	16	12.5	69037	2	AC014977	AC014977 Drosophi	c 614	16	12.5	110000	2	AC120911 3
542	16	12.5	69517	2	AC036117	AC036117 Homo sapi	c 615	16	12.5	110000	2	AC127928 2
543	16	12.5	70935	9	HS64115	AL049549 Human DNA	c 616	16	12.5	110000	2	AC127958 2
544	16	12.5	71618	8	AB012240	AB012240 Arabidops	c 617	16	12.5	110000	2	AC128494 0
545	16	12.5	71774	8	NCB9K17	BX294016 Neurospor	c 618	16	12.5	110000	2	AC128588 1
546	16	12.5	72593	9	AL441928	AL441928 Human DNA	c 619	16	12.5	110000	2	AC130865 2
547	16	12.5	72758	2	AC069108	AC069108 Homo sapi	c 620	16	12.5	110000	2	AC138832 3
548	16	12.5	74315	2	AC023029	AC023029 Homo sapi	c 621	16	12.5	110000	2	AC140288 2
549	16	12.5	74638	2	AC110343 4	Continuation (5 of	c 622	16	12.5	110000	2	AL928982 5
550	16	12.5	75044	2	AC125156 3	Continuation (4 of	c 623	16	12.5	110000	8	AC145127 08
551	16	12.5	76859	2	AC140004	AC140004 Rattus no	c 624	16	12.5	110275	10	AL663028
552	16	12.5	77344	9	AL645665	AL645665 Human DNA	c 625	16	12.5	110283	9	AC093245
553	16	12.5	77378	2	AC016183	AC016183 Homo sapi	c 626	16	12.5	110811	8	AP004520
554	16	12.5	78165	2	AC109967 3	Continuation (4 of	c 627	16	12.5	110819	9	AC055716
555	16	12.5	79554	8	NCB11B23	AL672287 Neurospor	c 628	16	12.5	112345	2	AP004049
556	16	12.5	81776	10	AL672287	AL672287 Mouse DNA	c 629	16	12.5	112695	8	AC119418
557	16	12.5	83701	2	AC095870 4	Continuation (5 of	c 630	16	12.5	112695	8	AC006476
558	16	12.5	85041	2	AC094428 5	Continuation (6 of	c 631	16	12.5	113687	2	AC104282
559	16	12.5	85565	9	AL136133	AL136133 Human DNA	c 632	16	12.5	113988	9	AL133546
560	16	12.5	86476	8	NCB8B20	AL355933 Neurospor	c 633	16	12.5	114653	9	AL590290
561	16	12.5	87503	9	AP004135	AP004135 Oryza sat	c 634	16	12.5	115296	2	AC084086
562	16	12.5	88110	9	AC108470	AC108470 Homo sapi	c 635	16	12.5	116696	3	PFMAL3P3
563	16	12.5	89016	9	HS86C11	AL021807 Human DNA	c 636	16	12.5	117270	2	AC005316
564	16	12.5	89379	2	AC140042	AC140042 Mus muscu	c 637	16	12.5	118447	2	AC093508
565	16	12.5	91945	2	AL353737	AL353737 Homo sapi	c 638	16	12.5	118677	5	BX255905
566	16	12.5	92625	2	AC006598	AC006598 Homo sapi	c 639	16	12.5	120607	2	AC122163
567	16	12.5	92800	9	AC004486	AC004486 Homo sapi	c 640	16	12.5	120647	2	AC122163
568	16	12.5	93458	9	AC0079344	AC0079344 Homo sapi	c 641	16	12.5	121442	2	AC132532
569	16	12.5	95087	10	AL806510	AL806510 Mouse DNA	c 642	16	12.5	122823	3	AC005996
570	16	12.5	95481	2	AC024192	AC024192 Homo sapi	c 643	16	12.5	123256	2	AL356602
571	16	12.5	95985	2	AP001344	AP001344 Homo sapi	c 644	16	12.5	123294	10	AL928633
572	16	12.5	96908	10	AC006542	AC006542 Mus muscu	c 645	16	12.5	123820	2	AC121236
573	16	12.5	96961	9	AC116563	AC116563 Homo sapi	c 646	16	12.5	123854	2	BX323827
574	16	12.5	97078	9	AC005841	AC005841 Homo sapi	c 647	16	12.5	124144	10	AL929270
575	16	12.5	99392	8	AC000103	AC000103 Genomic s	c 648	16	12.5	124786	9	AC115285
576	16	12.5	100000	9	AP000035	AP000035 Homo sapi	c 649	16	12.5	125032	9	AC007099

C 650	16	12.5 125590	3	CRY49E10	Z98866 Caenorhabdi	C 723	16	12.5 151156	2	EX465183	EX465183 Danio rer
C 651	16	12.5 125681	9	AL390736	Human DNA	724	16	12.5 151706	2	AC092744	AC092744 Homo sapi
C 652	16	12.5 125816	10	AL591127	Human DNA	725	16	12.5 151875	2	AC102003	AC102003 Mus muscu
C 653	16	12.5 126038	8	AP000367	Oryza sat	726	16	12.5 152306	9	AC113617	AC113617 Homo sapi
C 654	16	12.5 126106	10	AC125051	Mus muscu	C 727	16	12.5 152306	9	AC093596	AC093596 Homo sapi
C 655	16	12.5 126286	2	AC134954	Takifugu	C 728	16	12.5 152445	9	AC138410	AC138410 Homo sapi
C 656	16	12.5 126692	9	AL450997	Human DNA	C 729	16	12.5 152487	9	AP000431	AP000431 Homo sapi
C 657	16	12.5 127145	9	HS1119A7	Human DNA	C 730	16	12.5 152675	2	AC144556	AC144556 Homo sapi
C 658	16	12.5 127241	10	AL732551	Mouse DNA	C 731	16	12.5 152782	2	AL133255	AL133255 Human DNA
C 659	16	12.5 127811	2	AC008375	Homo sapi	C 732	16	12.5 152961	2	AC025021	AC025021 Homo sapi
C 660	16	12.5 128942	9	HS15B14	Human DNA	C 733	16	12.5 153300	2	AC073387	AC073387 Homo sapi
C 661	16	12.5 129218	2	AC138851	Homo sapi	C 734	16	12.5 153359	2	AC025000	AC025000 Homo sapi
C 662	16	12.5 129420	8	AC098694	Oryza sat	C 735	16	12.5 154062	2	AC118898	AC118898 Rattus no
C 663	16	12.5 129984	2	AC115620	Homo sapi	C 736	16	12.5 154235	9	HS428A13	HS428A13 Human DNA
C 664	16	12.5 129986	2	AC109151	Mus muscu	C 737	16	12.5 154361	2	AC084189	AC084189 Homo sapi
C 665	16	12.5 130043	8	AC126019	Medicago	C 738	16	12.5 154471	9	AC022523	AC022523 Homo sapi
C 666	16	12.5 130393	2	AP003565	Oryza sat	C 739	16	12.5 154502	2	EX511304	EX511304 Danio rer
C 667	16	12.5 131054	2	AC142062	Rattus no	C 740	16	12.5 154508	2	AC083974	AC083974 Homo sapi
C 668	16	12.5 131594	2	DMR17J10	Drosophila	C 741	16	12.5 154954	9	AC022507	AC022507 Homo sapi
C 669	16	12.5 131599	8	AC079685	Oryza sat	742	16	12.5 155731	9	AP001205	AP001205 Homo sapi
C 670	16	12.5 131972	9	AC008794	Homo sapi	743	16	12.5 155895	2	AC102166	AC102166 Mus muscu
C 671	16	12.5 132117	9	AC092463	Homo sapi	744	16	12.5 156221	2	AC134963	AC134963 Canis fam
C 672	16	12.5 132288	9	AC099848	Homo sapi	C 745	16	12.5 156820	9	AC074323	AC074323 Homo sapi
C 673	16	12.5 132638	9	AC026713	Homo sapi	C 746	16	12.5 156959	2	AC136042	AC136042 Rattus no
C 674	16	12.5 133413	8	OSJN00130	Oryza sat	747	16	12.5 156997	2	AF303734	AF303734 Homo sapi
C 675	16	12.5 133475	9	AC006254	Homo sapi	748	16	12.5 157069	9	AC010163	AC010163 Homo sapi
C 676	16	12.5 133924	4	AC093712	Canis fam	C 749	16	12.5 157266	9	AL353763	AL353763 Human DNA
C 677	16	12.5 135189	2	AP004851	Oryza sat	C 750	16	12.5 157381	2	AC129818	AC129818 Rattus no
C 678	16	12.5 135679	2	AC138993	Homo sapi	C 751	16	12.5 158043	9	AC105270	AC105270 Homo sapi
C 679	16	12.5 135758	9	AL136324	Human DNA	C 752	16	12.5 158390	9	AC138942	AC138942 Homo sapi
C 680	16	12.5 135924	8	AL115529	Rattus no	C 753	16	12.5 158412	2	AP001134	AP001134 Homo sapi
C 681	16	12.5 136263	8	AP003201	Oryza sat	C 754	16	12.5 158588	2	AC092273	AC092273 Homo sapi
C 682	16	12.5 136884	2	AC013397	Homo sapi	C 755	16	12.5 158591	9	AC012597	AC012597 Homo sapi
C 683	16	12.5 136884	2	AC013397	Homo sapi	C 756	16	12.5 158591	9	AC012597	AC012597 Homo sapi
C 684	16	12.5 137001	2	AC079088	Homo sapi	C 757	16	12.5 159050	2	AC109154	AC109154 Mus muscu
C 685	16	12.5 137769	9	AC002090	Genomic s	C 758	16	12.5 159098	2	AC016484	AC016484 Homo sapi
C 686	16	12.5 137995	2	AC025009	Homo sapi	759	16	12.5 159184	2	AC139484	AC139484 Homo sapi
C 687	16	12.5 139071	2	AC084275	Homo sapi	760	16	12.5 159468	2	AC140888	AC140888 Homo sapi
C 688	16	12.5 140031	2	AC121686	Rattus no	C 761	16	12.5 159775	10	AC131787	AC131787 Mus muscu
C 689	16	12.5 140478	2	AC138988	Homo sapi	762	16	12.5 160114	2	AC022181	AC022181 Homo sapi
C 690	16	12.5 140486	2	AC121624	Homo sapi	C 763	16	12.5 160115	9	AL161447	AL161447 Human DNA
C 691	16	12.5 141655	2	AC114894	Felis cat	C 764	16	12.5 160517	2	AL589785	AL589785 Homo sapi
C 692	16	12.5 141991	2	HSJ921D9	Homo sapi	765	16	12.5 160517	9	AC021032	AC021032 Homo sapi
C 693	16	12.5 142028	2	AC019575	Homo sapi	C 766	16	12.5 160635	9	AC118002	AC118002 Homo sapi
C 694	16	12.5 142723	9	HS1039K5	Human DNA	C 767	16	12.5 160741	9	AC097709	AC097709 Homo sapi
C 695	16	12.5 142745	9	AC124311	Homo sapi	C 768	16	12.5 161366	10	AC122505	AC122505 Mus muscu
C 696	16	12.5 144015	2	AL357144	Homo sapi	C 769	16	12.5 161460	9	AC093243	AC093243 Homo sapi
C 697	16	12.5 144117	8	AP003210	Oryza sat	C 770	16	12.5 161460	9	AL157879	AL157879 Human DNA
C 698	16	12.5 144404	8	AP003210	Oryza sat	C 771	16	12.5 161516	2	EX322538	EX322538 Danio rer
C 699	16	12.5 144869	2	AC114884	Rattus no	772	16	12.5 161549	9	AC044913	AC044913 Homo sapi
C 700	16	12.5 145021	2	AC093257	Homo sapi	773	16	12.5 161787	10	AL645484	AL645484 Mouse DNA
C 701	16	12.5 145493	2	AC138839	Homo sapi	C 774	16	12.5 161791	2	AC021690	AC021690 Homo sapi
C 702	16	12.5 145887	9	AC091435	Homo sapi	C 775	16	12.5 161893	2	AC137781	AC137781 Homo sapi
C 703	16	12.5 146582	10	AL662920	Mouse DNA	C 776	16	12.5 161921	2	AC133224	AC133224 Rattus no
C 704	16	12.5 146734	8	OSJN00179	Human DNA	777	16	12.5 162044	9	HSJ971B4	HSJ971B4 Human DNA
C 705	16	12.5 147111	10	AC138996	Homo sapi	C 778	16	12.5 162213	9	AC131071	AC131071 Mus muscu
C 706	16	12.5 147318	2	AC004738	Homo sapi	C 779	16	12.5 162255	2	AC026274	AC026274 Homo sapi
C 707	16	12.5 147404	9	AC004738	Homo sapi	C 780	16	12.5 162359	2	AC025274	AC025274 Homo sapi
C 708	16	12.5 147840	2	AC138939	Homo sapi	C 781	16	12.5 162533	2	AC119875	AC119875 Mus muscu
C 709	16	12.5 148081	2	AC068861	Homo sapi	C 782	16	12.5 162727	2	AC133202	AC133202 Mus muscu
C 710	16	12.5 148081	9	AB045361	Homo sapi	C 783	16	12.5 163273	2	AC142486	AC142486 Rattus no
C 711	16	12.5 148230	2	AC141548	Rattus no	784	16	12.5 163356	2	AC103362	AC103362 Mus muscu
C 712	16	12.5 148390	10	AC121930	Mus muscu	785	16	12.5 163558	2	AC142257	AC142257 Mus muscu
C 713	16	12.5 148448	2	AC120429	Mus muscu	786	16	12.5 163717	2	AC101953	AC101953 Mus muscu
C 714	16	12.5 148489	2	AL773528	Homo sapi	787	16	12.5 163899	2	AC123994	AC123994 Mus muscu
C 715	16	12.5 148489	2	AL773528	Homo sapi	C 788	16	12.5 164055	2	AC130883	AC130883 Rattus no
C 716	16	12.5 148930	9	AC069200	Homo sapi	C 789	16	12.5 164190	9	AC025448	AC025448 Homo sapi
C 717	16	12.5 149101	2	AC105982	Mus muscu	790	16	12.5 164704	2	AC016206	AC016206 Homo sapi
C 718	16	12.5 149235	2	AC068266	Homo sapi	C 791	16	12.5 164812	9	AC078820	AC078820 Homo sapi
C 719	16	12.5 150172	9	AC034227	Homo sapi	C 792	16	12.5 164964	9	AC015722	AC015722 Homo sapi
C 720	16	12.5 150188	2	AC019096	Homo sapi	C 793	16	12.5 165011	9	AL452555	AL452555 Human DNA
C 721	16	12.5 150481	9	HSJ961404	Human DNA	794	16	12.5 165152	9	AC012308	AC012308 Homo sapi
C 722	16	12.5 151030	10	AL805936	Mouse DNA	C 795	16	12.5 165180	2	AC119239	AC119239 Mus muscu

796	16	12.5 165351	9	AC019109	AC019109 Homo sapi	869	16	12.5 174367	3	AC023696	AC023696 Drosophil
797	16	12.5 165849	2	AC118597	AC118597 Mus muscu	870	16	12.5 174376	2	AC107366	AC107366 Mus muscu
c 798	16	12.5 165881	2	AC133732	AC133732 Rattus no	871	16	12.5 174548	2	AP002342	AP002342 Homo sapi
799	16	12.5 166258	9	AL805961	AL805961 Human DNA	c 872	16	12.5 174765	2	AC073103	AC073103 Homo sapi
c 800	16	12.5 166285	2	AC102421	AC102421 Mus muscu	c 873	16	12.5 174765	2	AL591385	AL591385 Human DNA
c 801	16	12.5 166321	2	AC018434	AC018434 Homo sapi	c 874	16	12.5 174805	2	AC145070	AC145070 Mus muscu
802	16	12.5 166924	8	MIATGENB	Y08502 A.thaliana	c 875	16	12.5 174852	2	AC010195	AC010195 Homo sapi
803	16	12.5 166937	8	AC025679	AC025679 Homo sapi	c 876	16	12.5 175036	9	AC097361	AC097361 Homo sapi
c 804	16	12.5 167281	2	AP001147	AP001147 Homo sapi	c 877	16	12.5 175306	9	AC079948	AC079948 Homo sapi
c 805	16	12.5 167305	2	AL590409	AL590409 Homo sapi	c 878	16	12.5 175321	2	AC129238	AC129238 Rattus no
c 806	16	12.5 167547	9	AC112673	AC112673 Homo sapi	c 879	16	12.5 175445	2	EX296546	EX296546 Danio rer
c 807	16	12.5 167691	9	AC117473	AC117473 Homo sapi	880	16	12.5 175470	9	AC113174	AC113174 Homo sapi
c 808	16	12.5 167721	9	AL158841	AL158841 Human DNA	881	16	12.5 175531	9	HSXKSRFXM	AL112157 Homo sapi
c 809	16	12.5 167849	2	AC119247	AC119247 Mus muscu	882	16	12.5 175632	2	AC115834	AC115834 Mus muscu
c 810	16	12.5 167908	9	AC122177	AC122177 Homo sapi	c 883	16	12.5 176048	9	AC025881	AC025881 Homo sapi
811	16	12.5 167912	9	AC060232	AC060232 Homo sapi	c 884	16	12.5 176152	2	AC135573	AC135573 Rattus no
c 812	16	12.5 167925	2	AC126993	AC126993 Rattus no	c 885	16	12.5 176172	2	AC127930	AC127930 Rattus no
c 813	16	12.5 168084	9	AC007037	AC007037 Homo sapi	c 886	16	12.5 176253	2	AP001078	AP001078 Homo sapi
814	16	12.5 168324	9	AL133391	AL133391 Human DNA	c 887	16	12.5 176320	9	AP003028	AP003028 Homo sapi
c 815	16	12.5 168419	2	AC025384	AC025384 Homo sapi	c 888	16	12.5 176417	2	AC024160	AC024160 Homo sapi
c 816	16	12.5 168547	2	AC138947	AC138947 Homo sapi	889	16	12.5 176444	2	AC132649	AC132649 Rattus no
c 817	16	12.5 168595	9	AC013401	AC013401 Homo sapi	c 890	16	12.5 176625	9	AC015977	AC015977 Homo sapi
c 818	16	12.5 168622	2	EX005050	EX005050 Danio rer	891	16	12.5 176630	2	AC010684	AC010684 Homo sapi
819	16	12.5 168717	2	AC024683	AC024683 Homo sapi	c 892	16	12.5 176760	9	AC018714	AC018714 Homo sapi
c 820	16	12.5 168741	2	AC013573	AC013573 Homo sapi	893	16	12.5 177071	2	AC128090	AC128090 Rattus no
c 821	16	12.5 168872	9	AC073325	AC073325 Homo sapi	894	16	12.5 177334	9	AC138069	AC138069 Homo sapi
c 822	16	12.5 168987	9	AC117390	AC117390 Homo sapi	c 895	16	12.5 177346	3	AC008231	AC008231 Drosophil
c 823	16	12.5 169043	2	AC018428	AC018428 Homo sapi	c 896	16	12.5 177414	2	EX247876	EX247876 Danio rer
c 824	16	12.5 169082	2	AC138933	AC138933 Homo sapi	c 897	16	12.5 177448	9	AC044781	AC044781 Homo sapi
c 825	16	12.5 169164	2	AC134747	AC134747 Rattus no	898	16	12.5 177463	3	AC126919	AC126919 Bos tauru
826	16	12.5 169177	2	AC023397	AC023397 Homo sapi	899	16	12.5 177720	9	AC007514	AC007514 Homo sapi
c 827	16	12.5 169226	9	AL590669	AL590669 Human DNA	c 900	16	12.5 177728	9	AL162425	AL162425 Human DNA
828	16	12.5 169246	2	AC091815	AC091815 Homo sapi	c 901	16	12.5 177947	2	AC074022	AC074022 Homo sapi
c 829	16	12.5 169252	2	AC126435	AC126435 Mus muscu	c 902	16	12.5 178057	10	AC125097	AC125097 Mus muscu
c 830	16	12.5 169261	9	AP000379	AP000379 Homo sapi	903	16	12.5 178360	2	EX530089	EX530089 Danio rer
c 831	16	12.5 169318	9	AL592425	AL592425 Homo sapi	904	16	12.5 178563	2	AP002419	AP002419 Homo sapi
c 832	16	12.5 169399	9	AC061965	AC061965 Homo sapi	c 905	16	12.5 178696	2	AC136892	AC136892 Gallus ga
c 833	16	12.5 169726	2	AC119920	AC119920 Mus muscu	c 906	16	12.5 178828	9	AC135707	AC135707 Human DNA
c 834	16	12.5 170000	2	AC004524	AC004524 Homo sapi	907	16	12.5 178857	2	AL114881	AL114881 Rattus no
c 835	16	12.5 170028	2	AP002435	AP002435 Homo sapi	c 908	16	12.5 178959	2	AC128118	AC128118 Rattus no
c 836	16	12.5 170066	9	AL161730	AL161730 Human DNA	909	16	12.5 178964	2	AC073155	AC073155 Homo sapi
837	16	12.5 170184	3	AC009911	AC009911 Drosophil	c 910	16	12.5 179203	9	AC009194	AC009194 Homo sapi
c 838	16	12.5 170371	9	AP001132	AP001132 Homo sapi	911	16	12.5 179253	5	AL928870	AL928870 Zebrafish
c 839	16	12.5 170560	2	AC127009	AC127009 Rattus no	c 912	16	12.5 179531	9	AC003080	AC003080 Homo sapi
840	16	12.5 170612	9	AC005317	AC005317 Homo sapi	c 913	16	12.5 179536	2	AC010814	AC010814 Homo sapi
c 841	16	12.5 170676	9	AC026258	AC026258 Homo sapi	c 914	16	12.5 179810	9	AC124997	AC124997 Homo sapi
c 842	16	12.5 170759	2	AC138994	AC138994 Homo sapi	c 915	16	12.5 179897	2	AC100735	AC100735 Mus muscu
c 843	16	12.5 170823	2	AC068173	AC068173 Mus muscu	c 916	16	12.5 180069	2	AC017041	AC017041 Homo sapi
c 844	16	12.5 170917	9	AC012560	AC012560 Homo sapi	917	16	12.5 180101	2	EX255923	EX255923 Homo sapi
c 845	16	12.5 170996	2	AC010938	AC010938 Homo sapi	c 918	16	12.5 180136	2	AC102758	AC102758 Mus muscu
c 846	16	12.5 171123	2	AC142009	AC142009 Rattus no	c 919	16	12.5 180526	2	AC078866	AC078866 Homo sapi
847	16	12.5 171452	2	AC122297	AC122297 Mus muscu	920	16	12.5 180919	10	AC127275	AC127275 Mus muscu
c 848	16	12.5 171452	9	AP000476	AP000476 Homo sapi	c 921	16	12.5 180999	9	AL513205	AL513205 Human DNA
c 849	16	12.5 171503	2	AC102190	AC102190 Mus muscu	c 922	16	12.5 181079	9	AC064802	AC064802 Homo sapi
c 850	16	12.5 171621	9	AP003177	AP003177 Homo sapi	c 923	16	12.5 181453	2	AC111426	AC111426 Rattus no
c 851	16	12.5 171660	2	AC017103	AC017103 Homo sapi	924	16	12.5 181548	2	AC141561	AC141561 Mus muscu
c 852	16	12.5 172281	9	AC068492	AC068492 Homo sapi	c 925	16	12.5 181603	3	AC124571	AC124571 Mus muscu
c 853	16	12.5 172418	10	AL627134	AL627134 Mouse DNA	926	16	12.5 181782	2	AC116970	AC116970 Rattus no
c 854	16	12.5 172505	9	AC124068	AC124068 Homo sapi	927	16	12.5 181854	10	AL806535	AL806535 Mouse DNA
c 855	16	12.5 172605	2	AC087753	AC087753 Homo sapi	928	16	12.5 181917	2	AC132728	AC132728 Rattus no
c 856	16	12.5 172624	2	AC127439	AC127439 Rattus no	929	16	12.5 182020	9	AP228660	AP228660 Homo sapi
857	16	12.5 172865	9	AC090791	AC090791 Homo sapi	c 930	16	12.5 182041	9	AC090921	AC090921 Homo sapi
c 858	16	12.5 172942	2	AC016899	AC016899 Homo sapi	c 931	16	12.5 182158	2	AC132161	AC132161 Rattus no
c 859	16	12.5 172953	9	AC067721	AC067721 Homo sapi	c 932	16	12.5 182212	5	AL928952	AL928952 Zebrafish
c 860	16	12.5 172991	9	AL353608	AL353608 Human DNA	933	16	12.5 182505	2	AC117789	AC117789 Mus muscu
861	16	12.5 173014	2	AC132070	AC132070 Atelerix	c 934	16	12.5 182673	2	AC073105	AC073105 Homo sapi
c 862	16	12.5 173174	2	AC141901	AC141901 Rattus no	c 935	16	12.5 182935	2	AC142131	AC142131 Rattus no
c 863	16	12.5 173415	9	CNS01DVS	AL136018 Human chr	c 936	16	12.5 183010	2	AC135952	AC135952 Macaca mu
c 864	16	12.5 173556	2	AC074388	AC074388 Homo sapi	937	16	12.5 183125	9	AP002848	AP002848 Homo sapi
c 865	16	12.5 173611	9	AC087280	AC087280 Homo sapi	c 938	16	12.5 183187	10	AC104328	AC104328 Mus muscu
c 866	16	12.5 173660	9	AC063979	AC063979 Homo sapi	c 939	16	12.5 183241	2	AC012111	AC012111 Homo sapi
c 867	16	12.5 173669	2	AC017110	AC017110 Homo sapi	c 940	16	12.5 183402	2	AC126231	AC126231 Bos tauru
c 868	16	12.5 173929	2	AC131928	AC131928 Homo sapi	c 941	16	12.5 183496	2	AC125075	AC125075 Mus muscu

942 12.5 183595 9 AC117460 Homo sapi  
943 12.5 183748 10 AL772351 Mouse DNA  
944 12.5 183947 2 AC138734 Papio anu  
945 12.5 184157 2 AC024728  
946 12.5 184497 2 AC099688 Homo sapi  
947 12.5 184661 2 AC134188  
948 12.5 184794 2 AP000849 Homo sapi  
949 12.5 184967 2 AC133299 Rattus no  
950 12.5 185163 2 AC133282 Mus muscu  
951 12.5 185203 2 AL161649 Homo sapi  
952 12.5 185272 9 AC105443 Homo sapi  
953 12.5 185405 3 AC023743 Drosophila  
954 12.5 185530 2 AL928859 Mus muscu  
955 12.5 185598 2 AC132909 Mus muscu  
956 12.5 185666 2 AC048389 Homo sapi  
957 12.5 185697 2 AC067814 Homo sapi  
958 12.5 186185 9 AC008817 Homo sapi  
959 12.5 186256 10 AL845333 Mouse DNA  
960 12.5 186361 2 AC079159 Homo sapi  
961 12.5 186659 2 AC127780 Rattus no  
962 12.5 186945 2 AC111427 Rattus no  
963 12.5 187072 9 AC099541 Homo sapi  
964 12.5 187164 2 AC129815 Rattus no  
965 12.5 187239 2 AC021349 Homo sapi  
966 12.5 187481 2 AC116817 Mus muscu  
967 12.5 187488 2 AC116394 Mus muscu  
968 12.5 187740 9 AC138409 Homo sapi  
969 12.5 187755 2 AC006718 Caenorhab  
970 12.5 187879 2 AC118479 Mus muscu  
971 12.5 187987 2 AC009639 Homo sapi  
972 12.5 188179 2 AC139286 Homo sapi  
973 12.5 188352 2 AC139496 Homo sapi  
974 12.5 188443 10 AC121775 Mus muscu  
975 12.5 188514 2 AC104554 Mus muscu  
976 12.5 188536 2 BX223873 Danio rer  
977 12.5 188593 2 AC136108 Rattus no  
978 12.5 188809 2 BX510943 Danio rer  
979 12.5 188829 9 AC112721 Homo sapi  
980 12.5 189145 2 AC126528 Rattus no  
981 12.5 189230 9 AC103974 Homo sapi  
982 12.5 189310 2 AC013419 Homo sapi  
983 12.5 189381 2 AC126920 Bos tauru  
984 12.5 189486 9 AC008733 Homo sapi  
985 12.5 189649 2 AC139641 Rattus no  
986 12.5 189652 2 BX537119 Danio rer  
987 12.5 189819 2 BX323462 Danio rer  
988 12.5 189949 9 AC005040 Homo sapi  
989 12.5 190294 2 AC138913 Homo sapi  
990 12.5 190641 10 AL626775 Mouse DNA  
991 12.5 190798 10 AC121811 Mus muscu  
992 12.5 190909 2 BX005369 Danio rer  
993 12.5 190987 2 AC113883 Rattus no  
994 12.5 190999 2 AC103899 Bos tauru  
995 12.5 191011 2 AC036138 Homo sapi  
996 12.5 191115 9 AC131053 Homo sapi  
997 12.5 191301 2 AC127820 Rattus no  
998 12.5 191544 2 AC138855 Homo sapi  
999 12.5 192400 2 AC022712 Homo sapi  
c1000 12.5 192753 2 AC107236 Mus muscu

## ALIGNMENTS

RESULT 1  
AF234618 4229 bp DNA linear PRI 29-MAR-2002  
LOCUS Homo sapiens megin gene, promoter region and partial sequence.  
DEFINITION AF234618  
ACCESSION AF234618  
VERSION AF234618.2 GI:19808130  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 4229)  
AUTHORS Miyata, T., Nangaku, M., Inagi, R. and Kurokawa, K.  
TITLE Transcriptional regulation of a mesangium-predominant gene, megin  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 4229)  
AUTHORS Miyata, T., Nangaku, M., Inagi, R. and Kurokawa, K.  
TITLE Direct Submission  
JOURNAL Submitted (15-FEB-2000) Institute of Medical Sciences and Department of Internal Medicine, Tokai University School of Medicine, Bohseidai, Isehara, Kanagawa 259-1193, Japan  
REFERENCE 3 (bases 1 to 4229)  
AUTHORS Miyata, T., Nangaku, M., Inagi, R. and Kurokawa, K.  
TITLE Direct Submission  
JOURNAL Submitted (29-MAR-2002) Institute of Medical Sciences and Department of Internal Medicine, Tokai University School of Medicine, Bohseidai, Isehara, Kanagawa 259-1193, Japan  
REMARK Sequence update by submitter  
COMMENT On Mar 29, 2002 this sequence version replaced gi:18000453.  
FEATURES  
Location/Qualifiers  
source  
1..4229  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18q21.3"  
promoter  
1..4021  
4022..>4229  
/product="megsin"  
BASE COUNT 1392 a 722 c 827 g 1288 t  
ORIGIN  
Query Match 41.4%; Score 53; DB 9; Length 4229;  
Best Local Similarity 100.0%; Pred. No. 3.2e-20;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 76 CTTGAGAGACAGTGTCTGCTCTGAGTCATAGGAGCCATCCAGAGCCAG 128  
DB 3996 CTTGAGAGACAGTGTCTGCTCTGAGTCATAGGAGCCATCCAGAGCCAG 4048  
AC072051 157284 bp DNA linear PRI 03-DEC-2001  
LOCUS Homo sapiens chromosome, clone RP11-79D21, complete sequence.  
DEFINITION AC072051  
ACCESSION AC072051  
VERSION AC072051.8 GI:16974280  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 157284)  
AUTHORS Birren, B., Linton, L., Nussbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome, clone RP11-79D21  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 157284)  
AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castie, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehotsky, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Minova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,

Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severi, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

# TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Submitted (07-JUN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severi, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

# TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Submitted (25-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 157284)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severi, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

# TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Submitted (03-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Nov 18, 2001 this sequence version replaced gi:14277307.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RX/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L966

Center clone name: 79\_D\_21

# FEATURES

## Source

1. 157284

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="RP11-79D21"

/clone\_lib="RPC1-11 Human Male BAC"

599..635

/rpt\_family="(T)n"

1040..1077

/rpt\_family="AT rich"

1411..1747

/rpt\_family="L1P"

2632..2992

/rpt\_family="MLT1A1"

3115..3138

/rpt\_family="AT rich"

complement(3487..3828)

/rpt\_family="L1MC4a"

3834..4338

/rpt\_family="MLT1C"

complement(4339..5397)

/rpt\_family="L1MC4a"

complement(5742..5899)

/rpt\_family="MER5B"

6499..6816

/rpt\_family="HAL1"

complement(7416..7621)

/rpt\_family="MER3"

complement(8163..9090)

/rpt\_family="L2"

9276..9419

/rpt\_family="MIR"

9598..9658

/rpt\_family="MER81"

9659..10033

/rpt\_family="MLT1J"

13986..14009

/rpt\_family="AT rich"

14011..14115

/rpt\_family="(TA)n"

complement(14942..15044)

/rpt\_family="MIR3"

15156..15298

/rpt\_family="L2"

15592..15773

/rpt\_family="L1MC4"

15874..18310

/rpt\_family="L1PA16"

16373..16374

/note="1328 bp of bacterial transposon insertion in BAC excised between these 2 bp"

18311..18331

/rpt\_family="(TAAAA)n"

18332..19566

/rpt\_family="L1PA16"

19567..19587

/rpt\_family="(CA)n"

19588..20042

/rpt\_family="L1PA16"

20043..20064

/rpt\_family="AT rich"

complement(21259..21386)

/rpt\_family="MIR3"

22395..22420

/rpt\_family="AT rich"

complement(22788..22948)

/rpt\_family="FRAM"

complement(23843..24253)

/rpt\_family="MSTA"

complement(24678..25041)

/rpt\_family="MLT1A2"

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Direct Submission  
 Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 189092)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
 Cook, A., Cooke, P., DeArlleano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
 Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,  
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,  
 Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N.,  
 Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J.,  
 Meneses, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Neill, A., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
 Pollara, V., Roymond, C., Rettig, R., Ribback, M., Riley, R., Rise, C.,  
 Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,  
 Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,  
 Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, G.,  
 Vassilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
 Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (30-MAR-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jun 1, 2002 this sequence version replaced G111990731.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIER  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L8516  
 Center clone name: 317\_G\_1  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 85659: contig of 85659 bp in length  
 \* 85660 85759: gap of 100 bp  
 \* 85760 163014: contig of 77255 bp in length  
 \* 163015 163114: gap of 100 bp  
 \* 163115 189092: contig of 25978 bp in length.  
 Location/Qualifiers  
 1. 189092  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="18"  
 /map="18"  
 /clone="RP11-317G1"  
 /clone\_lib="RP11-317G1"  
 BASE COUNT 57107 a 35683 c 36719 g 59369 t 214 others  
 ORIGIN  
 Query Match 41.4%; Score 53; DB 2; Length 189092;  
 Best Local Similarity 100.0%; Pred. No. 2,5e-20;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 QY 76 CTTGAGACAGCTGTCTCTGTCATGAGTCATGAGGAGCCATCCAGAGCCAG 128



```

Db      66639 CTTGAGACACAGTGTGCTGTCTGAGTCATAGGGAAGCCATCCAGAACCCAG 66691

RESULT 4
AX346917/c
LOCUS      AX346917      6182 bp      DNA      linear      PAT 01-FEB-2002
DEFINITION Sequence 1988 from Patent WO0200928.
ACCESSION  AX346917
VERSION     AX346917.1  GI:18494803
KEYWORDS    synthetic construct
SOURCE      synthetic construct
ORGANISM    artificial sequences.
REFERENCE   1
AUTHORS     Olek A., Piepenbrock, C. and Berlin, K.
TITLE       Diagnosis of diseases associated with the immune system
JOURNAL     Patent: WO 0200928-A 1988 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES   Location/Qualifiers
            source
            1..6182
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /note="Chemically treated genomic DNA (Homo sapiens)"
BASE COUNT  1934 a 37 c 1059 g 3152 t
ORIGIN
Query Match      16.4%; Score 21; DB 6; Length 6182;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5  AACTACATACAAACACACCTTA 25
        |||||||
Db      1306  AACTACATACAAACACACCTTA 1286

RESULT 5
BX296551
LOCUS      BX296551      194622 bp      DNA      linear      HTG 02-JUN-2003
DEFINITION Danio rerio clone CH211-224L10, *** SEQUENCING IN PROGRESS ***
ACCESSION  BX296551
VERSION     BX296551.2  GI:30348619
KEYWORDS    HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Bükaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 194622)
AUTHORS     McLaren, S.
TITLE       Direct Submission
JOURNAL     Submitted (28-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On May 2, 2003 this sequence version replaced gi:29335452.
COMMENT     ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfsh-help@sanger.ac.uk
            ----- Project Information
            Center project name: zc224l10
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 194521 bases at least Q40
            Consensus quality: 194528 bases at least Q30
            Consensus quality: 194562 bases at least Q20
            Insert size: 194622; sum-of-contigs
            Insert size: 193822; 5.2% error; agarose-fp
            Quality coverage: 9.83x in Q20 bases; sum-of-contigs Quality
            coverage: 10.08x in Q20 bases; agarose-fp
            -----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 194622: contig of 194622 bp in length.
* Location/Qualifiers
* 1..194622
* /organism="Danio rerio"
* /mol_type="genomic DNA"
* /db_xref="taxon:7955"
* /clone="CH211-224L10"
* /clone_lib="CHORI-211"
* 1..194622
* /note="assembly fragment:00958"
BASE COUNT  62945 a 34348 c 34000 g 63329 t
ORIGIN
Query Match      16.4%; Score 21; DB 2; Length 194622;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      53  AAAACCTAAATGCTTATAAGA 73
        |||||||
Db      56161  AAAACCTAAATGCTTATAAGA 56181

RESULT 6
AF396436
LOCUS      AF396436      47577 bp      DNA      linear      INV 11-MAR-2003
DEFINITION Tetrahymena thermophila strain SB210 mitochondrion, complete
            genome.
ACCESSION  AF396436
VERSION     AF396436.1  GI:15011465
KEYWORDS    mitochondrion Tetrahymena thermophila
SOURCE      Tetrahymena thermophila
ORGANISM    Bükaryota; Alveolata; Ciliophora; Oligohymenophorea;
            Hymenostomatida; Tetrahymenina; Tetrahymena.
REFERENCE   1  (bases 1 to 47577)
AUTHORS     Brunk, C.F., Lee, L.C., Tran, A.B. and Li, J.
TITLE       Complete sequence of the mitochondrial genome of Tetrahymena
            thermophila and comparative methods for identifying highly
            divergent genes
JOURNAL     Nucleic Acids Res. 31 (6), 1673-1682 (2003)
MEDLINE     22514407
PUBMED      12626709
REFERENCE   2  (bases 1 to 47577)
AUTHORS     Brunk, C.F., Tran, A.B., Lee, L.C. and Li, J.
TITLE       Direct Submission
JOURNAL     Submitted (27-JUN-2001) Organismic Biology, Ecology and Evolution,
            University of California, Los Angeles, 621 Charles E. Young Drive
            South, Los Angeles, CA 90095, USA
FEATURES   Location/Qualifiers
            source
            1..47577
            /organism="Tetrahymena thermophila"
            /organella="mitochondrion"
            /mol_type="genomic DNA"
            /strain="SB210"
            /db_xref="taxon:5911"
            complement(1..53)
            /note="telomere repeat"
            complement(65..145)
            /gene="trny"
            complement(65..145)
            /gene="trny"
            /product="tRNA-Tyr"
            /anticodon="(pos:110..112,aa:Tyr)
            complement(161..2844)
            misc_feature

```

```

gene
/translation="MKEFFQALNGYIFTKDTEFINTKKKIFGMWDYTYKIYNRPFKY
SENLSYKDTWKKFKSILSTFPKSRSLKNRMDTNIFLDPLNTYRGRLHSGKGLPVR
GQRTWNAWSYVRSNLNLEKFKLEAKRLYGNMETSLSLTIYLAQOLNYIKWLOWKRE
WLOARSRLKMONHEANFKNIDINSWSKGYIDGDFDKKEITKKKQQAOKNFTLIGFE
PGTFLYLRPNNSLSAKRNKIKIVLNEENKVKILQKKKKVVLKKKPEKKKNSWE
"
complement (7023 . .7478)
/gene="rps3"
CDS
complement (7023 . .7478)
/gene="rps3"
/codon_start=1
/transl_table=4
/product="ribosomal protein S3"
/protein_id="AAK7565.1"
/db_xref="GI:15011470"
/translation="MGLKSLPMLNKSGLSMYWNHNDWSIKLYKKSLSPLFLNEVINH
FLNENLYICIMKIRPTDRLKPRGNKSNINIKKSMNRHFLYLGKLLFLUKYQGWV
LVLINYCSRNLKIYINYSKFAKFKIAKSRQGVTSYVYKMDKYKPKF"
complement (7483 . .7779)
/gene="rps19"
CDS
complement (7483 . .7779)
/gene="rps19"
/codon_start=1
/transl_table=4
/product="ribosomal protein S19"
/protein_id="AAK7566.1"
/db_xref="GI:15011471"
/translation="MGSWRWLYKISLSIKSIKISIGIKKIKKISIMKSSITPS
SUTDNYIHKGYFKRELKIDYHVGFKFGEFILTRKPNIFPKKSNKSNKPRR"
complement (7784 . .8572)
/gene="rpl2"
CDS
complement (7784 . .8572)
/gene="rpl2"
/codon_start=1
/transl_table=4
/product="ribosomal protein L2"
/protein_id="AAK7567.1"
/db_xref="GI:15011472"
/translation="MELKVKKPTLRLPKEDIFKMTPIKKNIPFIKKNKSGRST
KQIIVHRGANKNKHLSINVRNQLNLAIIISLYINKNKSCFGLIKKSNCLSY
IKLANGFIGNYTKSLDKPLNFSFNRLILGSFIILMLAPKSFISNLISFNKKSQY
AKSAGTFLSIKIKELNFIILKPLTGKTKYFGSDSKATILGRNSFLNKFTVTGAGI
NNKMKWSVRGVAMPVDPHGGRTKTNQPEMSPGWITKKK"
complement (8576 . .9049)
/gene="ymf74"
CDS
complement (8576 . .9049)
/gene="ymf74"
/note="CPGN"
/codon_start=1
/transl_table=4
/product="ymf74"
/protein_id="AAK7568.1"
/db_xref="GI:15011473"
/translation="MKPKGEFLIKKKNFKFKFKILLIKYYSIRKLNLFYKIKLN
NFKIKTLINNNYFNITNGLDKYDNTQNFELNTLSIKNYKKNLIISNNQLD
IKFKOKPFIIDNKYVNSLICDNLDFDFIFISILTNLSLIEFYKNILINLIKIN"
complement (9062 . .9550)
/gene="nad10"
CDS
complement (9062 . .9550)
/gene="nad10"
/EC_number="1.6.5.3"
/codon_start=1
/transl_table=4
/product="NADH dehydrogenase subunit 10"
/protein_id="AAK7569.1"
/db_xref="GI:15011474"
/translation="MRVDYIVTPINAVYSWARQGSFWPLTFGLACCAVEMHATVSR
YDFDFGVIFRATPRQADLIIVAGTLTNKMAPGLRRLYEOTLDPKVLSMGSCANGGY
YHYSVVRGNCNRIIPVDIVPGCPPTBALMFGVLQLOKSLYRQINEKEKKNKDL
TY"
complement (9567 . .9968)
/gene="rps12"

```

```

/note="copy number 1; includes rnl_a, trnL and rnl_b"
complement (161 . .438)
/gene="rnl_a.1"
rRNA
complement (161 . .438)
/gene="rnl_a.1"
/product="large subunit ribosomal RNA"
complement (448 . .528)
/gene="trnL.1"
tRNA
complement (448 . .528)
/gene="trnL.1"
/product="tRNA-Leu"
/anticodon=(pos:492..494,aa:Leu)
complement (531 . .2844)
/gene="rnl_b.1"
rRNA
complement (531 . .2844)
/gene="rnl_b.1"
/product="large subunit ribosomal RNA"
complement (3023 . .3325)
/gene="ymf57"
CDS
complement (3023 . .3325)
/gene="ymf57"
/note="CPGN"
/codon_start=1
/transl_table=4
/product="ymf57"
/protein_id="AAK7561.1"
/db_xref="GI:15011466"
/translation="MLKKNLIKFKFPFVQSGFYVDRIKFKSEMPIRNIFYSSIFP
GEKFWIELTKTIDSDTIPNNRNFNINLVSKFLQILTLILYLFPIIFILYI"
complement (3335 . .4675)
/gene="ymf66"
CDS
complement (3335 . .4675)
/gene="ymf66"
/note="CPGN"
/codon_start=1
/transl_table=4
/product="ymf66"
/protein_id="AAK7562.1"
/db_xref="GI:15011467"
/translation="MGRNVLPHVNDVDFVFTTPQEPSTFKSVKPLPSDILGLG
VEWVTTSSEVLADYKLYLWNGSVNLDESDFNQYNAISLTSVSQLPYAVILD
RYSVLVQFTPTDNDWFRMLHSGKSTALILVHPELSWHINGLNQFTFYGILIEFV
YFDKNPDMCIVLHTLHILLILFLITGFVTILFSGYGNPTEENTIDSYLAASGT
VFAEKIISIDYLGCLFALAYVGFVYVHGTWSMLSHVLLISCVSIIIMFLFILG
MPTLLYDQGIFFLAYIKGAGKYSISVAEMDFYACILAFVIRILAQWIRVILWVTF
ISLSHVSDPDTINSALIGSENODSMNENLNTGSMVYILTVLPGKFIWLYELHT
FPVVCQVAPFAIVFWLFLYTFPIIEKHEDFPKGRERKKLKLNLNKN"
complement (4937 . .6154)
/gene="ymf76"
CDS
complement (4937 . .6154)
/gene="ymf76"
/note="CPGN"
/codon_start=1
/transl_table=4
/product="ymf76"
/protein_id="AAK7563.1"
/db_xref="GI:15011468"
/translation="MFLVKKIKKLKIKKIKKFKLYINYNFNLNNFVHKKNFLT
LYNNNNPXLIFVIXINLSKTSNKKLFLPLNPKLNKKKPKYKYSKIKPIL
NILKSKYKINLLFFIKYFSVLRKQSRIFNLARVKSURSKRPFKKLKKKTKAKY
FQMFKKKGRHKHYKILINLIDFFIRNKRFPRLHLYDIRKKYFRLNNRNITYFY
KFKIKHNFKIRIKISLSIKDRVHFYELSLRNIAIKLYAFTLRNANLFTKSGP
IFLNGHQLNPFKAYKGDITELPFQKFTILKLRNKKKMSNMKTKKNWRTLKNK
VNPEQRRLISRSFNTLNFKTKLTKLFQDYVRTLSYCVVLDITNFKEDLTYNKKLIP
IYLLKLPNKKIIS"
complement (6189 . .7019)
/gene="rps13"
CDS
complement (6189 . .7019)
/gene="rps13"
/codon_start=1
/transl_table=4
/product="ribosomal protein S13"
/protein_id="AAK7564.1"

```

CDS complement(9567. .9968)  
/gene="rps12"  
/codon start=1  
/transl table=4  
/product="ribosomal protein S12"  
/protein\_id="AAK77570.1"  
/db\_xref="GI:15011475"  
/translation="MGSICINKKKGNLHKLKRLARILGINPQKGGVVMKARIYTPKPKN  
SARFPAKVIIVNKKRLTAHIFGINHRRHSSVIVRGGRDLPGVRYTCVRGVYDF  
SKVLNFRNRSRYGIPKPDDEKKLRRFRV"  
complement(10019. .10555)  
/gene="nad2"  
complement(10019. .10555)  
/gene="nad2"

gene

CDS

Query Match 15.6%; Score 20; DB 3; Length 47577;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 AAACCTGGTTCAAAACCTAA 61  
|||||  
Db 6358 AAACCTGGTTCAAAACCTAA 6377

RESULT 7  
G30073  
LOCUS human STS SHGC-36146, sequence tagged site.  
DEFINITION  
ACCESSION G30073  
VERSION G30073.1 GI:1593624  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 316)  
AUTHORS Myers,R.M.  
JOURNAL Unpublished (1996)  
COMMENT  
Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu

Primer A: ACCTAATGCTTTATAACAATGGA  
Primer B: TTATTTGAGATTCCTCAATGAGC  
STS size: 115  
PCR Profile:  
Initial incubation: 94 degrees C for 90 seconds  
Denaturation: 94 degrees C for 15 seconds  
Annealing: 62 degrees C for 23 seconds  
Polymerization: 72 degrees C for 30 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9600

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Taq Polymerase: 0.05 units/ul  
Total Vol: 10 ul

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 20 mM  
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from H88366  
-- Washington University/Merck EST sequence.

FEATURES  
source  
Location/Qualifiers  
1. .316  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="19"  
STS  
primer\_bind 28. .52  
primer\_bind complement(118. .142)  
BASE COUNT 144 a 45 c 52 g 75 t  
ORIGIN

Query Match 14.8%; Score 19; DB 11; Length 316;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 AAAACCTAAATGCTTATAA 71  
|||||  
Db 25 AAAACCTAAATGCTTATAA 43

RESULT 8  
AX337914  
LOCUS Sequence 8423 from Patent WO0194629.  
DEFINITION  
ACCESSION AX337914  
VERSION AX337914.1 GI:18128633  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,  
Horrigan,S., Soppet,D.R. and Weaver,Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 8423 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source  
Location/Qualifiers  
1. .482  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 180 a 79 c 93 g 123 t  
ORIGIN

Query Match 14.8%; Score 19; DB 6; Length 482;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 AAAACCTAAATGCTTATAA 71  
|||||  
Db 41 AAAACCTAAATGCTTATAA 59

RESULT 9  
US8727/c  
LOCUS  
DEFINITION Caenorhabditis elegans cosmid D1005, complete sequence.  
ACCESSION US8727  
VERSION US8727.1 GI:1326257  
KEYWORDS HTG.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 34372)  
AUTHORS Waterston,R.  
TITLE Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium  
JOURNAL Science 282 (5396), 2012-2018 (1998)  
MEDLINE 99069613

9851916  
 PUBMED  
 2 (bases 1 to 34372)  
 REFERENCE  
 Wohlmann,P. and Hawkins,J.  
 TITLE  
 The sequence of C. elegans cosmid D1005  
 JOURNAL  
 Unpublished (2001)  
 REFERENCE  
 3 (bases 1 to 34372)  
 WATERSTON,R.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (19-MAY-1996) Robert Waterston  
 4 (bases 1 to 34372)  
 REFERENCE  
 WATERSTON,R.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (23-JUN-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 5 (bases 1 to 34372)  
 REFERENCE  
 WATERSTON,R.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (03-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 6 (bases 1 to 34372)  
 REFERENCE  
 WATERSTON,R.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (22-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 7 (bases 1 to 34372)  
 REFERENCE  
 WATERSTON,R.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (18-JUN-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 8 (bases 1 to 34372)  
 REFERENCE  
 WATERSTON,R.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (19-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 Submitted by:

Genome Sequencing Center  
 Department of Genetics, Washington University  
 St. Louis, MO 63110, USA, and  
 Sanger Centre, Hinxton Hall  
 Cambridge CB10 1RQ, England  
 email: rwnematode.wustl.edu and jessesanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

For a graphical representation of this cosmid sequence and its analysis see:  
<http://www.wormbase.org/db/seq/sequence?name=D1005;class=Sequence>

#### NEIGHBORING COSMID INFORMATION

The 5' cosmid is W05H7, 200 bp overlap; the 3' cosmid is F09E10, 700 bp overlap. Actual start of this cosmid is at base position 1 of D1005; actual end is at 6710 of F09E10.

#### NOTES:

Coding sequences below are the result of integration and manual review of the following data : computer analysis using the program

Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yui Kohara ([http://www.ddbj.nig.ac.jp/c-elegans/html/CE\\_INDEX.html](http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html)) and The C. elegans ORFome cloning project (<http://worfdb.dfci.harvard.edu/>), similarity to other proteins from BlastX analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

#### FEATURES

1. 34372  
 Location/Qualifiers  
 /organism="Caenorhabditis elegans"  
 /mol\_type="genomic DNA"  
 /strain="Bristol N2"  
 /db\_xref="taxon:6239"  
 /chromosome="X"  
 /clone="D1005"  
 /gene="D1005.3"  
 complement(49..1276)  
 /gene="D1005.3"  
 complement(join(49..109,159..904,954..1116,1170..1276))  
 /standard\_name="D1005.3"  
 /codon\_start=1  
 /product="Hypothetical protein D1005.3"  
 /protein\_id="AAB00581.1"  
 /db\_xref="GI:1326258"  
 /db\_xref="WormBase:D1005.3"  
 /translation="MTSSFSHFSSLSLFLFLFIASHREHSKFRKLKKWSKL  
 NYSHQKQALQKPLHVLRIQSEVAGFTDMFNNGASTSSAGSPARSDITIAASIQQR  
 DRERPVDFMETELDLGYLVLDLVDVTFDPAELQKCNILYDGEHPYEQEL  
 NGYSHVAGYGVPCGDYDQGYKMCVEKAETPDGATKTRAVRPPVDDYQKE  
 YSESSDNDGSDVDSDFEPKSKTKSAGLENEKPEQTRAKYKLABEAEETVK  
 LARRNDVAVRKAKGLQDKAEHDKRRIAELEGLQSERDARRDDQDTLEQ  
 LLRNGPKMQRPQRHLENFK"  
 6368..9935  
 /gene="D1005.2"  
 join(6368..6433,6496..6641,6693..6831,7538..7734,  
 8170..8481,8528..8694,9574..9935)  
 /gene="D1005.2"  
 /standard\_name="D1005.2"  
 /note="contains similarity to Pfam domain PF01704  
 (UTP--glucose-1-phosphate uridylyltransferase); coded for  
 by the following C. elegans cDNAs: yk423b3.3, yk513h5.3,  
 yk668d7.3, yk423b3.5, yk513h5.5, yk668d7.5, B1174222,  
 CSTR013H5.1"  
 /codon\_start=1  
 /product="Hypothetical protein D1005.2"  
 /protein\_id="AAB00582.2"  
 /db\_xref="GI:16604135"  
 /db\_xref="WormBase:D1005.2"  
 /translation="MSNDQLSKLREFFRQDQSEKAKHDSKIFEVLYSQFLENQHC  
 IDWNSKTEKHQVTLKOLEPFRKFNILKGLAVIKGLGTTMGCSKRSILVEV  
 RGYTFDLAVLEHQKCEAHVDVFLYIMNSFTDEDTKYLAEGYGNVTKFVQSK  
 CRPLDAETKLPTEDENEDWDGAWCPGEGNIQSQNSQLQDLADGREIIPVSN  
 DNTGANTDLOIQLMLDKNVYIMECTPKTVQVKGSTLIDIGRMWHLEMPQVPAEN  
 LPDQSTVKFKI ENTNNIYVNVKAVKLLPDKSEIIVNKKTIIRSEVLQESIGGC  
 IKNFNALCVHVERKRPVKNLGDLSLRSTLCDDHSTFKYHNEHLGAPVLSLD  
 PSILYNSVDVLDKFPFLPMDNCSFAVVDVTFKKNVLSGKVTNGKTESFGVVPD  
 GTVLKDQRYIAE"  
 complement(10363..11668)  
 /gene="D1005.6"  
 complement(join(10363..10458,10537..10685,10756..10804,  
 11567..11668))  
 /gene="D1005.6"  
 /standard\_name="D1005.6"  
 /note="coded for by the following C. elegans cDNAs:  
 CERSH12P"  
 /codon\_start=1  
 /product="Hypothetical protein D1005.6"  
 /protein\_id="AAB00583.1"  
 /db\_xref="GI:1326260"

```

/db_xref=" WormBase:D1005.6"
/translation="MGVAMQOAGALVGAAPPPPEAAEAIQSIRKAKLKLKFLKL
KPYFTSVDMKPKVKKKFKDKIFENCFCKKKKQNFQFFNFKCFKSENLEWSE
GRVTWNGALYDYDOLVARIIPHYDIRL"
complement(18246..20167)
/gene="D1005.5"
complement(join(18246..18341,18351..18483,18990..19142,
19194..19447,19901..20025,20073..20167))
/gene="D1005.5"
/standard_name="D1005.5"
/codon_start=1
/product="Hypothetical protein D1005.5"
/protein_id="AA000584.1"
/db_xref="GI:1326261"
/db_xref="WormBase:D1005.5"
/translation="MLLFTILVWFPCINVSFRKATCNTTEITWIPDKSDYFESQ
VELQTSRTGNPDSMAKIVQBECKFEKNVSKRPAIKVGAHFLYKFKVIVSV
QLLRLDVRGLIKLIFELKANRLDEYFRTSIQILMGEKPKSILNPKFVFI
KRIPIENQILHVSARADATCYQRKRSRVIVINLKIPQVQNKVKCVSPNNGSL
YHVBTESLRTLRVDIKKCFENLTFCPHAAVRBINCTDENLRYKISYGRKA"
26820..31335
/gene="D1005.1"
join(26820..26858,26982..27383,27511..27968,28015..28483,
28848..29312,29572..30175,30222..30521,30952..31335)
/gene="D1005.1"
/standard_name="D1005.1"
/note="contains similarity to Pfam domains PF00549
(CoA-1lgase), PF02629 (CoA binding domain); coded for by
the following C. elegans cDNAs: YK133b7.5, YK171b5.5,
YK133b7.3, YK200c9.5, YK277c5.5, YK416d10.5, YK200c9.3,
YK622f9.3, YK668b12.3, CEESB58F, YK171b5.3, YK277c5.3,
YK416d10.3, YK500a8.5, YK504b2.5, YK614d7.3, YK614d7.5,
YK622f9.5, YK627g4.5, YK668b12.5, YK678d5.3, YK678d5.5"
/codon_start=1
/product="Hypothetical protein D1005.1"
/protein_id="AA000585.1"
/db_xref="GI:1326262"
/db_xref="WormBase:D1005.1"
/translation="MSAKAVSELSGKVLKYKFPSPGLLSAPHAFVHVRAGENFDBIAN
14.8%; Score 19; DB 3; Length 34372;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 ACTACTTTGAAACCTGGTT 51
|||||
Db 22891 ACTACTTTGAAACCTGGTT 22873

RESULT 10
AC026356/c
LOCUS
DEFINITION
Homo sapiens 12 BAC RP11-81714 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
AC026356
AC026356.29 GI:15789206
VERSION
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 37490)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alshrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burket,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,I., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.B., Ding,Y., Dinu,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,

```

Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, J., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsif, O., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozano, R. J., Lu, A., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, B., Massey, E., Massey, E., McLeod, M. P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Minor, Z., Mitchell, T., Mohabbat, K., Montgomery, K. T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogut, M., Okwoni, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, J., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoshkar, N., Sisson, I., Sodergren, E., Sonaker, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleszyk, R., Wooten, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Kucheralapati, R., Weinstein, G. and Gibbs, R.

Direct Submission  
2 (bases 1 to 37490)  
Worley, K.C.  
Direct Submission  
Submitted (22-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 37490)  
Worley, K.C.  
Direct Submission  
Submitted (27-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 37490)  
Worley, K.C.  
Direct Submission  
Submitted (02-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Sep 27, 2001 this sequence version replaced gi:15145501:  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:  
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.  
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.  
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT.

## FEATURES

source Location/Qualifiers  
 1. 37490  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-81714"

## misc\_feature

1. 1531  
 /note="overlaps bases 5470. 7000 of clone AC087316"

## repeat\_region

/function="clone overlap"

## repeat\_region

/rpt\_family="AluJb"

## repeat\_region

/complement(3002..3250)

## repeat\_region

/rpt\_family="AluSc"

## repeat\_region

/complement(3285..3576)

## repeat\_region

/rpt\_family="AluSg"

## repeat\_region

/complement(3580..3878)

## repeat\_region

/rpt\_family="AluSg"

## repeat\_region

/complement(3986..4297)

## repeat\_region

/rpt\_family="AluSx"

## repeat\_region

4321..4351

## repeat\_region

/rpt\_family="AT rich"

## repeat\_region

/complement(5007..5300)

## repeat\_region

/rpt\_family="AluJb"

## repeat\_region

6722..6777

## STS

/rpt\_family="TA)n"

## repeat\_region

6877..6956

## repeat\_region

/standard\_name="86996"

## repeat\_region

/complement(7844..8143)

## misc\_feature

/rpt\_family="AluSg"

## repeat\_region

7849

## repeat\_region

/note="C ( T )"

## repeat\_region

/function="polymorphic site"

## repeat\_region

8251..8561

## repeat\_region

/rpt\_family="AluSp"

## repeat\_region

8674..8802

## repeat\_region

/rpt\_family="MER33"

## repeat\_region

8803..8843

## repeat\_region

/rpt\_family="TG)n"

## repeat\_region

8843..8979

## repeat\_region

/rpt\_family="TA)n"

## repeat\_region

8844..8979

## repeat\_region

/rpt\_family="MER33"

## repeat\_region

/complement(8980..9253)

## repeat\_region

/rpt\_family="AluJb"

## repeat\_region

9254..9332

## repeat\_region

/rpt\_family="MER33"

## repeat\_region

/complement(9333..9633)

## repeat\_region

/rpt\_family="AluJb"

## repeat\_region

9634..9724

## repeat\_region

/rpt\_family="MER33"

## repeat\_region

9961..10024

## repeat\_region

/rpt\_family="TATATG)n"

## repeat\_region

/complement(10200..10475)

## repeat\_region

/rpt\_family="AluJo"

## repeat\_region

11398..11501

## repeat\_region

/rpt\_family="L2"

repeat\_region 11811..11886  
 /rpt\_family="GA-rich"  
 repeat\_region 12127..12191  
 /rpt\_family="GA-rich"  
 repeat\_region 12951..13252  
 /rpt\_family="AluY"  
 repeat\_region 13263..13386  
 /rpt\_family="(TA)n"  
 repeat\_region complement(14022..14307)  
 /rpt\_family="AluJb"  
 STS 14789..15061  
 /standard\_name="171015"  
 STS 14792..14884  
 /standard\_name="50858"  
 repeat\_region 15422..15690  
 /rpt\_family="MER102"  
 repeat\_region 15894..15924  
 /rpt\_family="(GAAA)n"  
 STS 16938..17092  
 /standard\_name="50762"  
 repeat\_region 18200..18225  
 /rpt\_family="AT rich"  
 repeat\_region complement(18236..18539)  
 /rpt\_family="AluSx"  
 repeat\_region complement(18545..18671)  
 /rpt\_family="FLAM\_C"  
 repeat\_region 19365..19386  
 /rpt\_family="AT rich"  
 repeat\_region complement(19387..19479)  
 /rpt\_family="Alu"  
 repeat\_region 19618..19920  
 /rpt\_family="AluJb"

Query Match 14.8%; Score 19; DB 9; Length 37490;  
 Best Local Similarity 100.0%; Pred.No. 4.6;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AACTACATACACACCT 23  
 |||||

Db 6239 AACTACATACACACCT 6221

## RESULT 11

AP002543

LOCUS

AP002543 74350 bp DNA linear PLN 27-DEC-2000  
 Arabidopsis thaliana genomic DNA, chromosome 5, BAC clone: F15M7.

DEFINITION

AP002543 BAC00015

ACCESSION

AP002543.1 GI:8698605

VERSION

AP002543.1

KEYWORDS

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (sites)

AUTHORS

Kaneko,T., Katoh,T., Asamizu,E., Sato,S., Nakamura,Y., Kotani,H.

and Tabata,S.

STRUCTURAL ANALYSIS OF ARABIDOPSIS THALIANA CHROMOSOME 5. XI

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 74350)

AUTHORS

Nakamura,Y.

TITLE

Direct Submission

JOURNAL

Submitted (22-JUN-2000) Yasukazu Nakamura, Kazusa DNA Research

## COMMENT

Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)  
 Address for correspondence: kaos@kazusa.or.jp  
 For the latest information on annotation of this clone, please see [http://www.kazusa.or.jp/kaos/cgi-bin/agd\\_graph.cgi?c=F15M7](http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=F15M7)  
 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.  
 The software programs used to predict genes include: Grail

```

(Informatics Group, Oak Ridge National Laboratory,
http://combio.ornl.gov/Graill-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremlini.zozi.lastate.edu/cgi-bin/sp.cgi).
Genes encoding rRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MHF15 and the 3' clone is MPH15.
Location/Qualifiers
1. .74350
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="F15M7"
/clone_lib="IGF BAC"
complement(join(822..929,1126..1286,1468..1569,
1744..1834))
/notes="contains similarity to unknown protein
gb|AAF23194.1
gene_id:F15M7.1"
/codon_start=1
/evidence=not experimental
/protein_id="BAB11397.1"
/db_xref="GI:10178104"
/translation="MERIFAILPAAVFLILLVSPVIAVDVHYCEENAEYVKKYVD
ISPNPAPGPAFTTASANTGREISFGKLVIEVSFGVHSHETDLCTETSCPIQNG
DPLVAHSGVLPGVTPGSGYLLAKMKLDAKKELCIKFSFDIGLRASVADI"
complement(2599..3192)
/notes="gene_id:F15M7.2"
/codon_start=1
/evidence=not experimental
/product="C3HC4-type RING zinc finger protein-like"
/protein_id="BAB11398.1"
/db_xref="GI:10178105"
/translation="MARIVPPYSGHWTNTDRMGGLAYGIVSGILMLTTLTSL
YCTKSHISAPPTTRTRRQRENSGTLPPGCRFDDESDTVVVEVLGTEEVI
KGFKLPIYEARSYSIQKESSTSCSICLADYKQMDMLRVLPDCHLFDHNCVDPEW
LRLHPTCPVCTSPSPAMTPVADVPFSPRRPMDI"
7047..7775
/notes="contains similarity to MADS-box transcription
factor
gene_id:F15M7.3"
/codon_start=1
/evidence=not experimental
/protein_id="BAB11399.1"
/db_xref="GI:10178106"
/translation="MARKVRAWRDRMRASLRRLTGLIKKNELSLCDMRAS
VVFNRBEQTLAWPSPRANSLIDNFYSLTDHRTKAVDPESYQTVIEKIKGA
DTRKYTFEDELMTQVQNGRELADLPTRADKLIPIADKLMMLSKRMSGTGDAL
RASNVASGNGGLNMMETGRSFYVQKWFVDPQVQPCVDVETPLPTMVSGLDANNE
PSDELGTGKGSSNAGGAEADDA"
join(10042..10234,10407..10498,10594..10800,10877..11194)
/notes="gene_id:F15M7.4"
/codon_start=1
/evidence=not experimental
/product="transcription factor-like protein"
/protein_id="BAB11400.1"
/db_xref="GI:10178107"
/translation="MQTEELSPQTPWNNAFGSQPLTTESLSGEASQSFTGVKAVTT
EAEQGVNDKSTTLTFPSGKSRSDVPRPHVAFAMQACFFGFGAPMMYTKGPH
VQYQGVVSAVSQSRSGRWMLPLKMEETEDGTVVNSKQYHGIIRROSRKAKELSL
RCKPYMHSHSLHAMRPRGSGRFLNKTADAOKSPNSQSSEVFHPENETINS
SREANESLSDSATNDYFLSSAYSFGGVMPDKKNAAMDIGCCCKLNI"
complement(join(11853..11975,12059..13975))
/notes="gb|AAF63169.1
gene_id:F15M7.5

```

```

similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB11401.1"
/db_xref="GI:10178108"
/translation="MITSWHDLSRQNNQNVGADSESTRLVNIRNIAKIVNSRFR
SLSPDTPATHDAQVILWPSLDMPLLPVNFVSHLPIGHTFRKELELILKITAQFVAV
YKSTQRELVTRVSPSLFEFLKPTDGRNSFYTRIILGQVIMPSQKLKTSKPSFPG
PSKLIQAPVVKDDVEMAMTDLHAYVFANMINDIRAPAPVFAFTRGVTTTRNE
LDITIKIQAQVAVYGLFRTLMKRVFISPKFDFKSTDKSGFYLRVLDVGSVLR
SRKNGAGLGVVGFILKLDQVVEKDAVEMALTDLHALEFFANVDGVLOPRPEOVH
PMWTEPPLIMPQLSQSTFQVQPSCSPVPMSPPRPNDLQSGQSNKAPASV
APIPEPPEISCVENTALIVSKNGLETERKQSMELSMNDARHFWSTQVIFAFQOLK
AYERAQNDRAHDIQNVLRSGFPGFFPEKTEITLKELGIIKLTAQVARYGMNFVQ
LRKRVGNPQFLESTNNRSFYNGVIAYSRVLMPSKLSKSDDTATVVDGFLS
CLOLEKEGVIDMDLIDCFARLEADYSANVPOPHLSOTOMPTPTPTSPPLFL
PLKNESPVLADODHPSDIDEMQVIDFTQSLPDNVLASLADKTAGGIQIPANK"
complement(join(14937..15080,15164..15448,15539..15826,
15918..16226,16305..16488,16567..16722,16810..17045,
17454..17555))
/notes="gene_id:F15M7.6"
/codon_start=1
/evidence=not experimental
/product="ABC transporter-like protein"
/protein_id="BAB11402.1"
/db_xref="GI:10178109"
/translation="MNQCNVTVAEDIEAGKKKPKFOAETPLIFLFRDVTYKVI
KLTSSVEKEILITIGSGVNGEVIAMGPSGSKTLLSLAGRISOSSTGSGVTYN
DXPSYKLSKIGFTVQDDVFLPHLYKEVILTYAARLRLPKTLREQKKQALADIQE
LGLERCOTMIGGAFVGVSGERKRVISNEIIPSLDLDEPTSGLDSTALRIL
LMHIDNAAEKTIVTTIHQPSRSLFRHFDKLIILGSGSLLYFGKSEALDYFSSIGCS
PLIANNPABEQLLANGINDISVSELDROVGVNSGRETQTKRPSAAVHEYLVEA
YETVAQEKSKLLDPVLDREAKAKSTRKQWGTQWEOYCIILFCRGLKRRHEVF
SLWAVTOVLSAVILGLLWQSDIRTPMGLQDQAGLLEFIAVFWGFVFETAIAPFQ
ERAMNKERAADMYRLSAYFLARTISDLDFILPSLELLVYVETGLRISFPPEFLS
MLTVFLCLIAAQGLIAGLIMDLKATILASVTVMTFMAGGFVKASPFDFDLFC
F"
complement(23784..25652)
/notes="gene_id:F15M7.7"
/codon_start=1
/evidence=not experimental
/product="selenium-binding protein-like"
/protein_id="BAB11403.1"
/db_xref="GI:10178110"
/translation="MSNIVLNTLRKPKLALLQSCSSPSDLKIIGHFLRTHLSIDV
FVASRLIACYDSTENKFTLLGYAGIISOIQONNLFWNLJRCSTGAEKAF
GFTTOMLRKSWPNITFTFLKASSEMECVLGVGRQTHSQIVFGFQNDVYENSLVH
MYANGFIAAAGRFQMGFRDVSWSVAGYCKGVMENAREMPDENPHNLTWS
IMINGYAKNCFEKAIDLFEFMREGVAVANSTVMVSVISSCAHLGALBGERAYEVV
KSHMTVNLIIGTALVDMFWRCGDIKKAITHVPEGLPETDSLWSIIKGLAVGHAHKA
MHYFSOMISLGFTIPRDTFTAVLSACHGGLVEKGLIYENMKOHGIBPRLHVGCI
VDMGLRAGLIAEAENFILMWHVKNAPILGALLGACKIYKNTVAERVGNMLIKVPE
HSGYVLLSNYACAGQWDKTSLDMMKELVKYKPPGWSLEIIDGKINKFTMGDDQK
HPENGKLRKWSBELKILGIGTGNTOGDAFVDVEEKESSIMHSEKIALAYGMK
TKPOTTRIVKLRVCECHVTYKLISVYGRVRLVDRNRHFFHFGVCGCRDYW"
complement(26394..27671)
/notes="contains similarity to unknown protein
gb|AAF71807.1
gene_id:F15M7.8"
/codon_start=1
/evidence=not experimental
/protein_id="BAB11404.1"
/db_xref="GI:10178111"
/translation="MEKCNILLKSKRSKSKLKHQHPESLFPKEVEEDEDDE
GGFKLIAASQSHGVQVQLNLYFNPGAVNVRVYGLNQLLSDELVIDLGLGANH
LGVLIATYASQSHGVQVQLNLYFNPGAVNVRVYGLNQLLSDELVIDLGLGANH
SNLKIIDFYDYLQFQSWLCANLEMKPRKLRDNIITVRGIVSEYDFTFEPKFPVLL
EGCLDGLWPALEKKSRYLTKVGVDFEAVGFWMKLEKTFYRSDGARBERPLFDPK
FAEYKVLDSYDVPVYPRDLFGVIGNERDVRWIIIGPAGSSSFHIDPNSTSAWN
AVITGSKKVLFPFDPVPGVHPSDGAEVACPVSIENWFMPYDDTKDWEKKEIECI
CKAGEVMFVNGWWHLVINLEESIAITQNYASR"
complement(join(28281..28383,28556..28602,28740..30146))
/notes="gb|AAF23201.1

```

```

gene id:F15M7.9
strong similarity to unknown protein"
/codon start=1
/evidence=not experimental
/protein_id="BAB11405.1"
/db_xref="GI:10178112"
/translation="MDIAIVSSPTDVRVRCDCGCSLNGASFGSLRSVVKRYEFP
ENEKLFHPELEIDLSNAKVOIENELLETITVSSQOQSIODLYELDEERNAASTA
ASEAMSLILORDKAELOMELEQKFAEKMEHDOELLDELIVRECTOALT
FRAQAKHMMNSFGFTEAEVETHEKMLNSFMSIENDYQIDLTSTSYPIKCNVNEP
GLEADIDDDVEKYPADSPHPLKTLERRISQMERNPSTFTGDSVSGRRHYTEKNV
FQSGPRHGRFRTVSTGSSALLGTTREKRLDFGNDPRNNGSFRKEDPPEAAGNS
FARDGDSRTIGKMDNRVTYIDSVHSVSGTARQKFNADTDYAMSPRIISNQ
PLDGPETSLKLMLOALADRESMEQAIMSMRTKQAWLLKXIAOHLSDXVYPERE
LPLRSTLIIGAFNISVFWKITSFVFWKAKRSKYNGVQGNMGLQMLEKPRIR
QWRCLSSQV"
complement(join(32678..32984,34934..35616))
/note="gb|AAD04946.2
gene id:F15M7.10
similar to unknown protein"
/codon start=1
/evidence=not experimental
/protein_id="BAB11406.1"
/db_xref="GI:10178113"
/translation="WGSIGBEPQVARDCMGLLQLLSNGTVLRSESIDLITQIIPKNN

Query Match      14.8%; Score 19; DB 8; Length 74350;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      15 CAACCACTTAGTCAGATA 33
      |||||
Db      26951 CAACCACTTAGTCAGATA 26969

RESULT 12
AP001152
LOCUS      Homo sapiens genomic DNA, chromosome 11q, clone:RP11-856F16,
DEFINITION complete sequence.
ACCESSION AP001152
VERSION AP001152.4 GI:14517583
KEYWORDS HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Homo sapiens genomic DNA
JOURNAL      Published Only in Database (2000)
REFERENCE      2 (bases 1 to 91200)
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL      Submitted (08-FEB-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Sueno-choi,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(BE-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT      On Jun 20, 2001 this sequence version replaced gi:11094158.
FEATURES
SOURCE      1..91200
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11q"
            /clone="RP11-856F16"
BASE COUNT      25425 a 18733 c 20156 g 26886 t
ORIGIN

```

```

Query Match      14.8%; Score 19; DB 9; Length 91200;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      86 AGTGTGTGCTCTGAGTCA 104
      |||||
Db      60891 AGTGTGTGCTCTGAGTCA 60909

RESULT 13
AC087861/c
LOCUS      Felis catus clone RP86-49M22, complete sequence.
DEFINITION AC087861
ACCESSION AC087861.2 GI:15150612
VERSION AC087861.2
KEYWORDS HTG.
SOURCE      Felis catus (cat)
ORGANISM      Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE      1 (bases 1 to 107365)
AUTHORS      Akhtar,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-O.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Masiello,C., Mastrian,S.D.,
McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A., Shevchenko,Y.,
Stancipop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W.,
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Zhang,L.-H. and Green,E.D.
TITLE      NISC Comparative Sequencing Initiative
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 107365)
AUTHORS      Green,E.D.
TITLE      Direct Submission
JOURNAL      Submitted (02-FEB-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
REFERENCE      3 (bases 1 to 107365)
AUTHORS      Green,E.D.
TITLE      Direct Submission
JOURNAL      Submitted (14-AUG-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
REFERENCE      4 (bases 1 to 107365)
AUTHORS      Green,E.D.
TITLE      Direct Submission
JOURNAL      Submitted (03-NOV-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
COMMENT      On Aug 14, 2001 this sequence version replaced gi:12656795.
            ----- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC
            Web site: http://www.nisc.nih.gov
            Contact: nisc.mouse@nih.gov
            ----- Project Information
            Center project name: awt
            Center clone name: 049M22

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.
            Location/Qualifiers
            1..107365
            /organism="Felis catus"
FEATURES
SOURCE

```



```

/mol_type="genomic DNA"
/db_xref="taxon:9685"
/clone="RP86-49M22"
/clone_lib="RP86"
<1.20642
misc_feature
    /note="this sequence is not the entire insert of clone
    RP86-49M22; clone overlaps with GenBank Accession Number
    AC087807 (nucleotides 79244-121506) clone RP86-294B21
    (center project name awv); this annotated segment
    represents overlap with nucleotides 100872-121506 of
    AC087807"
misc_feature
    107321..107355
    /note="clone has a very small overlap with GenBank
    Accession Number AC087731 (nucleotides 1-43) clone
    RP86-45908 (center project name awv)"
BASE COUNT 32771 a 21013 c 19977 g 33604 t
ORIGIN
Query Match 14.8%; Score 19; DB 4; Length 107365;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 TAGTCAGACTACTTTGCA 42
|||||
Db 70771 TAGTCAGACTACTTTGCA 70753

RESULT 14
AC011007/c 147556 bp DNA linear HTG 04-MAY-2001
LOCUS Homo sapiens chromosome 11 clone RP11-61E17 map 11, WORKING DRAFT
DEFINITION
SEQUENCE, 5 unordered pieces.
AC011007
AC011007.3 GI:8567837
HTG: HTGS_P19E1; HTGS_DRAFT.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 147556)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,A., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Casle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hags,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lien,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (29-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 15, 2000 this sequence version replaced gi:7655500.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIPR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1354
Center clone name: 61_E_17

```

```

----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator-amersham; 3% of reads
Chemistry: Dye-terminator Big Dye; 97% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 141528 bases at least Q40
Consensus quality: 14522 bases at least Q30
Consensus quality: 14579 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 147156; sum-of-contigs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 base.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 6645: contig of 6645 bp in length
* 6646 6745: gap of 100 bp
* 6746 31131: contig of 24386 bp in length
* 31132 31231: gap of 100 bp
* 31232 60186: contig of 28955 bp in length
* 60187 60286: gap of 100 bp
* 60287 94486: contig of 34200 bp in length
* 94487 94586: gap of 100 bp
* 94587 147556: contig of 52970 bp in length.
FEATURES
    location/Qualifiers
    1..147556
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11"
    /clone="RP11-61E17"
    /clone_lib="RPC1-11 Human Male BAC"
    1..6645
    /note="assembly_fragment"
    6746..31131
    /note="assembly_fragment"
    31232..60186
    /note="assembly_fragment"
    clone_end:SP6
    vector_side:left"
    60287..94486
    /note="assembly_fragment"
    clone_end:T7
    vector_side:left"
    94587..147556
    /note="assembly_fragment"
BASE COUNT 42156 a 31237 c 31280 g 42481 t
ORIGIN
Query Match 14.8%; Score 19; DB 2; Length 147556;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 AGTCTGTCTCTGAGTCA 104
|||||
Db 57425 AGTCTGTCTCTGAGTCA 57407

RESULT 15
AC003682/c 153875 bp DNA linear PRI 29-JUN-1998
LOCUS Homo sapiens chromosome 19, overlapping cosmids F18547, F11133,
DEFINITION R27945, R28830 and R32804, complete sequence.
AC003682
AC003682.1 GI:3264845
VERSION HTG.
SOURCE Homo sapiens (human)

```

```

ORGANISM Homo sapiens
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 153875)
Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Adamson,A.W.,
Burkhardt-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stillwagen,S.,
Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Carnes,J.,
Danganan,L., Poundstone,P., Christensen,M., Georgescu,A., Avila,J.,
Liu,S., Attix,C., Andreise,T., Frankheim,M., Amico-Keller,G.,
Coefield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,
Kronmiller,B., Arellano,A., Montgomery,M., Ow,D., Nolan,M.,
Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.
Sequence analysis of a 2 Mb region containing a zinc finger (ZNF)
gene cluster in 19q13.4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 153875)
AUTHORS Lamerdin,J.E.
JOURNAL Direct Submission
TITLE Submitted (16-DEC-1997) Human Genome Center, Lawrence Livermore
JOURNAL National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
REFERENCE 3 (bases 1 to 153875)
AUTHORS Lamerdin,J.E.
JOURNAL Direct Submission
TITLE Submitted (29-JUN-1998) Joint Genome Institute, Lawrence Livermore
JOURNAL National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
COMMENT On Jun 29, 1998 this sequence version replaced gi:2689440.
Map and sequence oriented from g centromere to telomere.
This accession derived from cosmid F18547 from bases 1 to 37,736,
cosmid F1133 from bases 36,861 to 40,261, cosmid R27945 from bases
40,194 to 77,593, cosmid R28830 from bases 74,683 to 118,033, and
cosmid R32804 from bases 114,809 to 153,875. This accession
overlaps cosmid F25419 (AC003005) to the left from bases 1 to
5,157, and overlaps cosmid F20191 (AC004017) from bases 150,049 to
153,875. No sequence discrepancies were identified in any of the
overlapping regions between cosmids.

Additional chr 19 map and sequence information are available at:
http://www-bio.lnl.gov/bbrp/genome/genome.html.
FEATURES
source
1..153875
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="19"
/map="19q13.4 between D19S773 to ZNF134"
/clons="Splice of sequence from cosmids
F18547-F1133-R27945-R28830-R32804"
/cell_line="UV5HU9-5B for F library clones, 5HL2-B for R"
/notes="Cosmid libraries L19NC02 and L19NC03 were
constructed at LNL from flow-sorted chromosomes from
human-hamster hybrids UV5HL9-5B and 5HL2-B, respectively,"
which carry chromosome 19 as their only human chromosome."
complement(1..62)
/rpt_family="Alu"
repeat_region 101..248
/rpt_family="Alu"
repeat_region 247..808
/rpt_family="L1"
repeat_region 813..1114
/rpt_family="L1MB8"
repeat_region 1115..1325
/rpt_family="TIGGER2"
complement(1115..1325)
/rpt_family="AluSc"
repeat_region 1329..1752
/rpt_family="L1MB8"
repeat_region 2245..4769
/rpt_family="L1MB8"
repeat_region 6786..7215
/rpt_family="HERVK"
repeat_region 7217..7296
/rpt_family="L1MB8"
complement(7217..7296)
/rpt_family="Alu"
repeat_region 7518..7794
/rpt_family="Alu"
repeat_region 7795..7919
complement(7795..7919)
/rpt_family="AluJo"
repeat_region 7923..8719
complement(7923..8719)
/rpt_family="TIGGER2"
repeat_region 8723..9023
complement(8723..9023)
/rpt_family="AluSc"
repeat_region 9024..9489
complement(9024..9489)
/rpt_family="TIGGER2"
repeat_region 9498..9529
complement(9498..9529)
/rpt_family="MER28"
repeat_region 9530..9607
complement(9530..9607)
/rpt_family="L1MB8"
repeat_region 9612..10109
complement(9612..10109)
/rpt_family="Alu"
repeat_region 10110..10335
complement(10110..10335)
/rpt_family="AluSc"
repeat_region 10336..10637
complement(10336..10637)
/rpt_family="Alu"
repeat_region 10638..11019
complement(10638..11019)
/rpt_family="Alu"
repeat_region 11020..11319
complement(11020..11319)
/rpt_family="Alu"
repeat_region 11320..12161
complement(11320..12161)
/rpt_family="GGA"
repeat_region 12162..13598
complement(12162..13598)
/rpt_family="Alu"
repeat_region 13599..13912
complement(13599..13912)
/rpt_family="AluSc"
repeat_region 13913..14242
complement(13913..14242)
/rpt_family="AluSp"
repeat_region 14243..15288
complement(14243..15288)
/rpt_family="AluSc"
repeat_region 15289..15699
complement(15289..15699)
/rpt_family="MER4B"
repeat_region 15700..17123
complement(15700..17123)
/rpt_family="LINE2"
repeat_region 17124..17917
complement(17124..17917)
/rpt_family="MER7A"
repeat_region 17918..18164
complement(17918..18164)
/rpt_family="MER4A"
repeat_region 18165..18616
complement(18165..18616)
/rpt_family="Alu"
repeat_region 18617..18639
complement(18617..18639)
/rpt_family="TABA"
repeat_region 18640..18801
complement(18640..18801)
/rpt_family="MER4C"
repeat_region 18802..19438
complement(18802..19438)
/rpt_family="MER21B"
repeat_region 19439..19716
complement(19439..19716)
/rpt_family="MIR"
repeat_region 19717..19820
complement(19717..19820)
/rpt_family="MER4C"
repeat_region 19821..20668
complement(19821..20668)
/rpt_family="MER4C"
repeat_region 20669..20930
complement(20669..20930)
/rpt_family="L1"
repeat_region 20931..21180
complement(20931..21180)
/rpt_family="MER5A"
repeat_region 21181..21830
complement(21181..21830)
/rpt_family="AluSc"
repeat_region 21831..22641
complement(21831..22641)
/rpt_family="L1MB8"
repeat_region 22642..27639
complement(22642..27639)
/rpt_family="L1MB8"
repeat_region 27640..28561
complement(27640..28561)
/rpt_family="L1MB8"
repeat_region 28562..29788
complement(28562..29788)
/rpt_family="L1MB8"
repeat_region 29789..30668
complement(29789..30668)
/rpt_family="L1MB8"
repeat_region 30669..31519
complement(30669..31519)
/rpt_family="L1MB8"
repeat_region 31520..32471
complement(31520..32471)
/rpt_family="L1MB8"
repeat_region 32472..33423
complement(32472..33423)
/rpt_family="L1MB8"
repeat_region 33424..34375
complement(33424..34375)
/rpt_family="L1MB8"
repeat_region 34376..35327
complement(34376..35327)
/rpt_family="L1MB8"
repeat_region 35328..36279
complement(35328..36279)
/rpt_family="L1MB8"
repeat_region 36280..37231
complement(36280..37231)
/rpt_family="L1MB8"
repeat_region 37232..38183
complement(37232..38183)
/rpt_family="L1MB8"
repeat_region 38184..39135
complement(38184..39135)
/rpt_family="L1MB8"
repeat_region 39136..40087
complement(39136..40087)
/rpt_family="L1MB8"
repeat_region 40088..41039
complement(40088..41039)
/rpt_family="L1MB8"
repeat_region 41040..41991
complement(41040..41991)
/rpt_family="L1MB8"
repeat_region 41992..42943
complement(41992..42943)
/rpt_family="L1MB8"
repeat_region 42944..43895
complement(42944..43895)
/rpt_family="L1MB8"
repeat_region 43896..44847
complement(43896..44847)
/rpt_family="L1MB8"
repeat_region 44848..45799
complement(44848..45799)
/rpt_family="L1MB8"
repeat_region 45800..46751
complement(45800..46751)
/rpt_family="L1MB8"
repeat_region 46752..47703
complement(46752..47703)
/rpt_family="L1MB8"
repeat_region 47704..48655
complement(47704..48655)
/rpt_family="L1MB8"
repeat_region 48656..49607
complement(48656..49607)
/rpt_family="L1MB8"
repeat_region 49608..50559
complement(49608..50559)
/rpt_family="L1MB8"
repeat_region 50560..51511
complement(50560..51511)
/rpt_family="L1MB8"
repeat_region 51512..52463
complement(51512..52463)
/rpt_family="L1MB8"
repeat_region 52464..53415
complement(52464..53415)
/rpt_family="L1MB8"
repeat_region 53416..54367
complement(53416..54367)
/rpt_family="L1MB8"
repeat_region 54368..55319
complement(54368..55319)
/rpt_family="L1MB8"
repeat_region 55320..56271
complement(55320..56271)
/rpt_family="L1MB8"
repeat_region 56272..57223
complement(56272..57223)
/rpt_family="L1MB8"
repeat_region 57224..58175
complement(57224..58175)
/rpt_family="L1MB8"
repeat_region 58176..59127
complement(58176..59127)
/rpt_family="L1MB8"
repeat_region 59128..60079
complement(59128..60079)
/rpt_family="L1MB8"
repeat_region 60080..61031
complement(60080..61031)
/rpt_family="L1MB8"
repeat_region 61032..61983
complement(61032..61983)
/rpt_family="L1MB8"
repeat_region 61984..62935
complement(61984..62935)
/rpt_family="L1MB8"
repeat_region 62936..63887
complement(62936..63887)
/rpt_family="L1MB8"
repeat_region 63888..64839
complement(63888..64839)
/rpt_family="L1MB8"
repeat_region 64840..65791
complement(64840..65791)
/rpt_family="L1MB8"
repeat_region 65792..66743
complement(65792..66743)
/rpt_family="L1MB8"
repeat_region 66744..67695
complement(66744..67695)
/rpt_family="L1MB8"
repeat_region 67696..68647
complement(67696..68647)
/rpt_family="L1MB8"
repeat_region 68648..69599
complement(68648..69599)
/rpt_family="L1MB8"
repeat_region 69600..70551
complement(69600..70551)
/rpt_family="L1MB8"
repeat_region 70552..71503
complement(70552..71503)
/rpt_family="L1MB8"
repeat_region 71504..72455
complement(71504..72455)
/rpt_family="L1MB8"
repeat_region 72456..73407
complement(72456..73407)
/rpt_family="L1MB8"
repeat_region 73408..74359
complement(73408..74359)
/rpt_family="L1MB8"
repeat_region 74360..75311
complement(74360..75311)
/rpt_family="L1MB8"
repeat_region 75312..76263
complement(75312..76263)
/rpt_family="L1MB8"
repeat_region 76264..77215
complement(76264..77215)
/rpt_family="L1MB8"
repeat_region 77216..78167
complement(77216..78167)
/rpt_family="L1MB8"
repeat_region 78168..79119
complement(78168..79119)
/rpt_family="L1MB8"
repeat_region 79120..80071
complement(79120..80071)
/rpt_family="L1MB8"
repeat_region 80072..81023
complement(80072..81023)
/rpt_family="L1MB8"
repeat_region 81024..81975
complement(81024..81975)
/rpt_family="L1MB8"
repeat_region 81976..82927
complement(81976..82927)
/rpt_family="L1MB8"
repeat_region 82928..83879
complement(82928..83879)
/rpt_family="L1MB8"
repeat_region 83880..84831
complement(83880..84831)
/rpt_family="L1MB8"
repeat_region 84832..85783
complement(84832..85783)
/rpt_family="L1MB8"
repeat_region 85784..86735
complement(85784..86735)
/rpt_family="L1MB8"
repeat_region 86736..87687
complement(86736..87687)
/rpt_family="L1MB8"
repeat_region 87688..88639
complement(87688..88639)
/rpt_family="L1MB8"
repeat_region 88640..89591
complement(88640..89591)
/rpt_family="L1MB8"
repeat_region 89592..90543
complement(89592..90543)
/rpt_family="L1MB8"
repeat_region 90544..91495
complement(90544..91495)
/rpt_family="L1MB8"
repeat_region 91496..92447
complement(91496..92447)
/rpt_family="L1MB8"
repeat_region 92448..93399
complement(92448..93399)
/rpt_family="L1MB8"
repeat_region 93400..94351
complement(93400..94351)
/rpt_family="L1MB8"
repeat_region 94352..95303
complement(94352..95303)
/rpt_family="L1MB8"
repeat_region 95304..96255
complement(95304..96255)
/rpt_family="L1MB8"
repeat_region 96256..97207
complement(96256..97207)
/rpt_family="L1MB8"
repeat_region 97208..98159
complement(97208..98159)
/rpt_family="L1MB8"
repeat_region 98160..99111
complement(98160..99111)
/rpt_family="L1MB8"
repeat_region 99112..100063
complement(99112..100063)
/rpt_family="L1MB8"
repeat_region 100064..101015
complement(100064..101015)
/rpt_family="L1MB8"
repeat_region 101016..101967
complement(101016..101967)
/rpt_family="L1MB8"
repeat_region 101968..102919
complement(101968..102919)
/rpt_family="L1MB8"
repeat_region 102920..103871
complement(102920..103871)
/rpt_family="L1MB8"
repeat_region 103872..104823
complement(103872..104823)
/rpt_family="L1MB8"
repeat_region 104824..105775
complement(104824..105775)
/rpt_family="L1MB8"
repeat_region 105776..106727
complement(105776..106727)
/rpt_family="L1MB8"
repeat_region 106728..107679
complement(106728..107679)
/rpt_family="L1MB8"
repeat_region 107680..108631
complement(107680..108631)
/rpt_family="L1MB8"
repeat_region 108632..109583
complement(108632..109583)
/rpt_family="L1MB8"
repeat_region 109584..110535
complement(109584..110535)
/rpt_family="L1MB8"
repeat_region 110536..111487
complement(110536..111487)
/rpt_family="L1MB8"
repeat_region 111488..112439
complement(111488..112439)
/rpt_family="L1MB8"
repeat_region 112440..113391
complement(112440..113391)
/rpt_family="L1MB8"
repeat_region 113392..114343
complement(113392..114343)
/rpt_family="L1MB8"
repeat_region 114344..115295
complement(114344..115295)
/rpt_family="L1MB8"
repeat_region 115296..116247
complement(115296..116247)
/rpt_family="L1MB8"
repeat_region 116248..117199
complement(116248..117199)
/rpt_family="L1MB8"
repeat_region 117200..118151
complement(117200..118151)
/rpt_family="L1MB8"
repeat_region 118152..119103
complement(118152..119103)
/rpt_family="L1MB8"
repeat_region 119104..120055
complement(119104..120055)
/rpt_family="L1MB8"
repeat_region 120056..121007
complement(120056..121007)
/rpt_family="L1MB8"
repeat_region 121008..121959
complement(121008..121959)
/rpt_family="L1MB8"
repeat_region 121960..122911
complement(121960..122911)
/rpt_family="L1MB8"
repeat_region 122912..123863
complement(122912..123863)
/rpt_family="L1MB8"
repeat_region 123864..124815
complement(123864..124815)
/rpt_family="L1MB8"
repeat_region 124816..125767
complement(124816..125767)
/rpt_family="L1MB8"
repeat_region 125768..126719
complement(125768..126719)
/rpt_family="L1MB8"
repeat_region 126720..127671
complement(126720..127671)
/rpt_family="L1MB8"
repeat_region 127672..128623
complement(127672..128623)
/rpt_family="L1MB8"
repeat_region 128624..129575
complement(128624..129575)
/rpt_family="L1MB8"
repeat_region 129576..130527
complement(129576..130527)
/rpt_family="L1MB8"
repeat_region 130528..131479
complement(130528..131479)
/rpt_family="L1MB8"
repeat_region 131480..132431
complement(131480..132431)
/rpt_family="L1MB8"
repeat_region 132432..133383
complement(132432..133383)
/rpt_family="L1MB8"
repeat_region 133384..134335
complement(133384..134335)
/rpt_family="L1MB8"
repeat_region 134336..135287
complement(134336..135287)
/rpt_family="L1MB8"
repeat_region 135288..136239
complement(135288..136239)
/rpt_family="L1MB8"
repeat_region 136240..137191
complement(136240..137191)
/rpt_family="L1MB8"
repeat_region 137192..138143
complement(137192..138143)
/rpt_family="L1MB8"
repeat_region 138144..139095
complement(138144..139095)
/rpt_family="L1MB8"
repeat_region 139096..140047
complement(139096..140047)
/rpt_family="L1MB8"
repeat_region 140048..141000
complement(140048..141000)
/rpt_family="L1MB8"
repeat_region 141001..141952
complement(141001..141952)
/rpt_family="L1MB8"
repeat_region 141953..142904
complement(141953..142904)
/rpt_family="L1MB8"
repeat_region 142905..143856
complement(142905..143856)
/rpt_family="L1MB8"
repeat_region 143857..144808
complement(143857..144808)
/rpt_family="L1MB8"
repeat_region 144809..145760
complement(144809..145760)
/rpt_family="L1MB8"
repeat_region 145761..146712
complement(145761..146712)
/rpt_family="L1MB8"
repeat_region 146713..147664
complement(146713..147664)
/rpt_family="L1MB8"
repeat_region 147665..148616
complement(147665..148616)
/rpt_family="L1MB8"
repeat_region 148617..149568
complement(148617..149568)
/rpt_family="L1MB8"
repeat_region 149569..150520
complement(149569..150520)
/rpt_family="L1MB8"
repeat_region 150521..151472
complement(150521..151472)
/rpt_family="L1MB8"
repeat_region 151473..152424
complement(151473..152424)
/rpt_family="L1MB8"
repeat_region 152425..153376
complement(152425..153376)
/rpt_family="L1MB8"
repeat_region 153377..154328
complement(153377..154328)
/rpt_family="L1MB8"
repeat_region 154329..155280
complement(154329..155280)
/rpt_family="L1MB8"
repeat_region 155281..156232
complement(155281..156232)
/rpt_family="L1MB8"
repeat_region 156233..157184
complement(156233..157184)
/rpt_family="L1MB8"
repeat_region 157185..158136
complement(157185..158136)
/rpt_family="L1MB8"
repeat_region 158137..159088
complement(158137..159088)
/rpt_family="L1MB8"
repeat_region 159089..160040
complement(159089..160040)
/rpt_family="L1MB8"
repeat_region 160041..160992
complement(160041..160992)
/rpt_family="L1MB8"
repeat_region 160993..161944
complement(160993..161944)
/rpt_family="L1MB8"
repeat_region 161945..162896
complement(161945..162896)
/rpt_family="L1MB8"
repeat_region 162897..163848
complement(162897..163848)
/rpt_family="L1MB8"
repeat_region 163849..164800
complement(163849..164800)
/rpt_family="L1MB8"
repeat_region 164801..165752
complement(164801..165752)
/rpt_family="L1MB8"
repeat_region 165753..166704
complement(165753..166704)
/rpt_family="L1MB8"
repeat_region 166705..167656
complement(166705..167656)
/rpt_family="L1MB8"
repeat_region 167657..168608
complement(167657..168608)
/rpt_family="L1MB8"
repeat_region 168609..169560
complement(168609..169560)
/rpt_family="L1MB8"
repeat_region 169561..170512
complement(169561..170512)
/rpt_family="L1MB8"
repeat_region 170513..171464
complement(170513..171464)
/rpt_family="L1MB8"
repeat_region 171465..172416
complement(171465..172416)
/rpt_family="L1MB8"
repeat_region 172417..173368
complement(172417..173368)
/rpt_family="L1MB8"
repeat_region 173369..174320
complement(173369..174320)
/rpt_family="L1MB8"
repeat_region 174321..175272
complement(174321..175272)
/rpt_family="L1MB8"
repeat_region 175273..176224
complement(175273..176224)
/rpt_family="L1MB8"
repeat_region 176225..177176
complement(176225..177176)
/rpt_family="L1MB8"
repeat_region 177177..178128
complement(177177..178128)
/rpt_family="L1MB8"
repeat_region 178129..179080
complement(178129..179080)
/rpt_family="L1MB8"
repeat_region 179081..180032
complement(179081..180032)
/rpt_family="L1MB8"
repeat_region 180033..180984
complement(180033..180984)
/rpt_family="L1MB8"
repeat_region 180985..181936
complement(180985..181936)
/rpt_family="L1MB8"
repeat_region 181937..182888
complement(181937..182888)
/rpt_family="L1MB8"
repeat_region 182889..183840
complement(182889..183840)
/rpt_family="L1MB8"
repeat_region 183841..184792
complement(183841..184792)
/rpt_family="L1MB8"
repeat_region 184793..185744
complement(184793..185744)
/rpt_family="L1MB8"
repeat_region 185745..186696
complement(185745..186696)
/rpt_family="L1MB8"
repeat_region 186697..187648
complement(186697..187648)
/rpt_family="L1MB8"
repeat_region 187649..188600
complement(187649..188600)
/rpt_family="L1MB8"
repeat_region 188601..189552
complement(188601..189552)
/rpt_family="L1MB8"
repeat_region 189553..190504
complement(189553..190504)
/rpt_family="L1MB8"
repeat_region 190505..191456
complement(190505..191456)
/rpt_family="L1MB8"
repeat_region 191457..192408
complement(191457..192408)
/rpt_family="L1MB8"
repeat_region 192409..193360
complement(192409..193360)
/rpt_family="L1MB8"
repeat_region 193361..194312
complement(193361..194312)
/rpt_family="L1MB8"
repeat_region 194313..195264
complement(194313..195264)
/rpt_family="L1MB8"
repeat_region 195265..196216
complement(195265..196216)
/rpt_family="L1MB8"
repeat_region 196217..197168
complement(196217..197168)
/rpt_family="L1MB8"
repeat_region 197169..198120
complement(197169..198120)
/rpt_family="L1MB8"
repeat_region 198121..199072
complement(198121..199072)
/rpt_family="L1MB8"
repeat_region 199073..200024
complement(199073..200024)
/rpt_family="L1MB8"
repeat_region 200025..200976
complement(200025..200976)
/rpt_family="L1MB8"
repeat_region 200977..201928
complement(200977..201928)
/rpt_family="L1MB8"
repeat_region 201929..202880
complement(201929..202880)
/rpt_family="L1MB8"
repeat_region 202881..203832
complement(202881..203832)
/rpt_family="L1MB8"
repeat_region 203833..204784
complement(203833..204784)
/rpt_family="L1MB8"
repeat_region 204785..205736
complement(204785..205736)
/rpt_family="L1MB8"
repeat_region 205737..206688
complement(205737..206688)
/rpt_family="L1MB8"
repeat_region 206689..207640
complement(206689..207640)
/rpt_family="L1MB8"
repeat_region 207641..208592
complement(207641..208592)
/rpt_family="L1MB8"
repeat_region 208593..209544
complement(208593..209544)
/rpt_family="L1MB8"
repeat_region 209545..210496
complement(209545..210496)
/rpt_family="L1MB8"
repeat_region 210497..211448
complement(210497..211448)
/rpt_family="L1MB8"
repeat_region 211449..212400
complement(211449..212400)
/rpt_family="L1MB8"
repeat_region 212401..213352
complement(212401..213352)
/rpt_family="L1MB8"
repeat_region 213353..214304
complement(213353..214304)
/rpt_family="L1MB8"
repeat_region 214305..215256
complement(214305..215256)
/rpt_family="L1MB8"
repeat_region 215257..216208
complement(215257..216208)
/rpt_family="L1MB8"
repeat_region 216209..217160
complement(216209..217160)
/rpt_family="L1MB8"
repeat_region 217161..218112
complement(217161..218112)
/rpt_family="L1MB8"
repeat_region 218113..219064
complement(218113..219064)
/rpt_family="L1MB8"
repeat_region 219065..220016
complement(219065..220016)
/rpt_family="L1MB8"
repeat_region 220017..220968
complement(220017..220968)
/rpt_family="L1MB8"
repeat_region 220969..221920
complement(220969..221920)
/rpt_family="L1MB8"
repeat_region 221921..222872
complement(221921..222872)
/rpt_family="L1MB8"
repeat_region 222873..223824
complement(222873..223824)
/rpt_family="L1MB8"
repeat_region 223825..224776
complement(223825..224776)
/rpt_family="L1MB8"
repeat_region 224777..225728
complement(224777..225728)
/rpt_family="L1MB8"
repeat_region 225729..226680
complement(225729..226680)
/rpt_family="L1MB8"
repeat_region 226681..227632
complement(226681..227632)
/rpt_family="L1MB8"
repeat_region 227633..228584
complement(227633..228584)
/rpt_family="L1MB8"
repeat_region 228585..229536
complement(228585..229536)
/rpt_family="L1MB8"
repeat_region 229537..230488
complement(229537..230488)
/rpt_family="L1MB8"
repeat_region 230489..231440
complement(230489..231440)
/rpt_family="L1MB8"
repeat_region 231441..232392
complement(231441..232392)
/rpt_family="L1MB8"
repeat_region 232393..233344
complement(232393..233344)
/rpt_family="L1MB8"
repeat_region 233345..234296
complement(233345..234296)
/rpt_family="L1MB8"
repeat_region 234297..235248
complement(234297..235248)
/rpt_family="L1MB8"
repeat_region 235249..236200
complement(235249..236200)
/rpt_family="L1MB8"
repeat_region 236201..237152
complement(236201..237152)
/rpt_family="L1MB8"
repeat_region 237153..238104
complement(237153..238104)
/rpt_family="L1MB8"
repeat_region 238105..239056
complement(238105..239056)
/rpt_family="L1MB8"
repeat_region 239057..240008
complement(239057..240008)
/rpt_family="L1MB8"
repeat_region 240009..240960
complement(240009..24096
```

```

FYKQSLDHRHHTGERPYECGCKGKAFIHKGLLEHQRHHTGKPYVVCICGKSF1
RSSYMHQRHHTGERPYECGCKGKAFIHKGLLEHQRHHTGKPYVVCICGKSF1
VKLLQHQHHTHTRHHTGERPYECGCKGKAFIHKGLLEHQRHHTGKPYVVCICGKSF1
LLOKQVHSGERPYNCTACEKAFIYKNLVEHQRHHTGKPYECGCKGKAFNKYSLV
RHQKHHTERP"
23856. .23978
/note="predicted exon, program: rail2exons_human_1.3,
frame: 2, quality: good, score: 72.000"
repeat_region
24902. .25017
/rpt_family="MER47B"
repeat_region
25019. .25213
/rpt_family="MER47B"
repeat_region
25519. .25814
/rpt_family="AluJb"
misc_feature
25855. .27639
/note="predicted exon, program: rail2exons_human_1.3,
frame: 0, quality: excellent, score: 84.000"
28460. .28498
/note="predicted exon, program: rail2exons_human_1.3,
frame: 2, quality: good, score: 53.000"
repeat_region
29190. .29486
/rpt_family="AluSx"
repeat_region
29556. .29954
/rpt_family="AluSx"
repeat_region
31008. .31058
/rpt_family="IIM4"
repeat_region
31339. .31368
/rpt_family="AT rich"
repeat_region
complement(31425..31613)
/rpt_family="MTIB"
misc_feature
complement(31821..31892)
/note="predicted exon, program: rail2exons_human_1.3,
frame: 0, quality: excellent, score: 83.000"
repeat_region
32367. .32667
/rpt_family="AluSq"
repeat_region
33552. .33847
/rpt_family="AluSx"
repeat_region
34313. .34436
/rpt_family="FLAM_C"
repeat_region
34443. .34735
/rpt_family="AluTo"
complement(join(35687..36705,44943..45069))
/note="Hypothetical Kruppel- Type Zinc Finger Protein;
Most similar to PID|d1024590 (AB007873) KIAA0413 [Homo
sapiens]"
/codon_start=1
/product="R27945_1"

```

```

Query Match      14.8%; Score 19; DB 9; Length 153875;
Best Local Similarity 100.0%; Pred.No. 4.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      53 AAAACCTAAAGCTTTATA 71
Db      132712 AAAACCTAAAGCTTTATA 132694

```

Search completed: August 14, 2003, 19:12:39  
Job time : 1755 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 17:28:28 ; Search time 177 Seconds  
(without alignments)  
1952.136 Million cell updates/sec

Title: US-09-889-611-1  
Perfect score: 128  
Sequence: 1 aatgaactacatacaacac.....gaagccatcccaagaagccag 128

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_19Jun03.\*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128	100.0	128	21	AAA71434 Human meglin promo
2	128	100.0	1431	21	AAA71435 Human meglin promo
3	30	23.4	30	21	AAA71448 Human meglin promo
4	30	23.4	30	21	AAA71449 Human meglin promo
5	30	23.4	30	21	AAA71450 Human meglin promo
6	25	19.5	25	21	AAA71453 Human meglin promo
7	22	17.2	23	21	AAA71452 Human meglin promo
8	21	16.4	6182	24	ABU34015 Human immune syste

9	19	14.8	482	24	ABL70086	Pancreas cancer re
c 10	19	14.8	1523	21	AAZ39949	Arabidopsis thalia
11	18	14.1	47	21	AAZ69316	Human map-related
c 12	18	14.1	1227	23	AAZ83451	DNA encoding novel
13	18	14.1	168174	24	ABT11173	Human 5-lipoxygena
c 14	18	14.1	168273	24	ABT11114	Human 5-lipoxygena
c 15	18	14.1	169998	24	AD36511	Human Hgr-1 gene.
c 16	18	14.1	197496	24	ABN85584	Human EGFR SEQ ID
17	17	13.3	24	22	AAH63146	Shrimp white spot
c 18	17	13.3	686	22	ABK41670	CDNA encoding nove
c 19	17	13.3	936	22	AAH52933	S. epidermidis ope
c 20	17	13.3	1032	24	ABN93184	Staphylococcus epi
c 21	17	13.3	1103	22	AAH62834	Shrimp white spot
c 22	17	13.3	1245	24	ABK84665	Human cDNA differe
c 23	17	13.3	1261	19	AAK14011	H. pylori GHPO 165
c 24	17	13.3	1397	21	AAAS2642	Eosinophil activat
c 25	17	13.3	1807	24	ABQ93314	Human cDNA SEQ ID
c 26	17	13.3	1846	24	ABQ93313	Human cDNA SEQ ID
c 27	17	13.3	3356	22	AAH54556	S. epidermidis gen
c 28	17	13.3	3376	22	AAH54081	S. epidermidis gen
c 29	17	13.3	9201	15	AAQ55003	Humanised anti-CD1
c 30	17	13.3	16918	24	ABL33617	Human immune syste
c 31	17	13.3	86574	24	ABK83560	Human cDNA differe
c 32	17	13.3	160271	22	AAH85750	Bipolar affective
c 33	17	13.3	160271	22	AAH85756	Human chromosome 1
c 34	17	13.3	160271	22	AAH85756	Human chromosome 1
c 35	17	13.3	160271	22	AAH85756	Human chromosome 1
c 36	17	13.3	160271	22	AAH85756	Human chromosome 1
c 37	17	13.3	160271	22	AAH85756	Human chromosome 1
c 38	17	13.3	160271	22	AAH85756	Human chromosome 1
c 39	17	13.3	160271	22	AAH85756	Human chromosome 1
c 40	17	13.3	305107	22	AAH62689	Nucleotide sequenc
c 41	16	12.5	60	24	ABN38463	Shrimp white spot
c 42	16	12.5	346	25	ABX42130	Human spliced tran
c 43	16	12.5	380	22	AAH65106	Bovine EGT associa
c 44	16	12.5	385	24	ABL64865	Novel human polynu
c 45	16	12.5	385	24	ABL64865	Lung cancer relate
c 46	16	12.5	475	23	ABV11926	Kidney cancer rela
c 47	16	12.5	505	24	ABT10503	Human prostate exp
c 48	16	12.5	594	25	ABZ54691	Human breast cance
c 49	16	12.5	735	4	AAH30161	Aspergillus oryzae
c 50	16	12.5	753	24	ABK96740	Sequence encoding
c 51	16	12.5	756	24	ABK96748	DNA encoding human
c 52	16	12.5	758	24	ABK96745	DNA encoding human
c 53	16	12.5	763	4	AAH30162	Sequence of the E.
c 54	16	12.5	763	11	AAQ06498	IFN-alpha 76 gene
c 55	16	12.5	763	19	AAV33298	Escherichia coli t
c 56	16	12.5	845	22	AAH22059	Human breast cance
c 57	16	12.5	850	22	AAH22059	Human breast cance
c 58	16	12.5	876	4	AAH30108	Sequence encoding
c 59	16	12.5	876	11	AAQ06497	IFN-alpha 76 gene
c 60	16	12.5	876	19	AAV33297	DNA encoding inter
c 61	16	12.5	896	24	ABK96746	DNA encoding human
c 62	16	12.5	958	20	AAH37501	Human secreted pro
c 63	16	12.5	977	6	AAH50357	Human interferon-a
c 64	16	12.5	1292	24	AAH519307	Mouse elongation o
c 65	16	12.5	1563	24	ABZ31933	Candida albicans e
c 66	16	12.5	1652	23	ABL07443	Drosophila melanog
c 67	16	12.5	1686	2	AAH10014	Sequence encoding
c 68	16	12.5	1758	11	AAQ06094	Reverse transcript
c 69	16	12.5	1920	21	AAH66792	Human membrane-ty
c 70	16	12.5	1938	22	AAH66792	Human DNA encoding
c 71	16	12.5	1938	22	AAH66792	Human MTSWMP polyn
c 72	16	12.5	1938	24	ABX10011	Human MTSWMP DNA f
c 73	16	12.5	1938	24	ABV78192	Human polynucleoti
c 74	16	12.5	1938	24	ABV78192	Human polynucleoti
c 75	16	12.5	2118	22	AAH28238	Nucleotide sequenc
c 76	16	12.5	2118	22	AAH28238	Human DNA encoding
c 77	16	12.5	2315	24	ABK89382	Gibberella fujikur
c 78	16	12.5	2423	12	AAQ11727	Sequence encoding
c 79	16	12.5	2423	16	AAQ91979	msDNA-Ec67 retron
c 80	16	12.5	2423	21	AAZ86865	E. coli msDNA-Ec67
c 81	16	12.5	2423	24	ABQ77596	Escherichia coli C

82	16	12.5	2595	22	ABAI18355	Human nervous syst	155	11.7	479	24	ABS03174	Human genome-deriv
83	16	12.5	2595	25	ABZ74021	Secreted protein g	156	11.7	500	20	AAV86067	EST clone D4. Hom
84	16	12.5	2595	25	ABZ76021	Human secreted pro	c 157	11.7	526	20	ABQ31568	Oligonucleotide fo
85	16	12.5	2620	21	AAAI4342	cDNA encoding huma	158	11.7	526	24	ABQ31569	Oligonucleotide fo
86	16	12.5	2620	21	AAAI13380	Human MTS-MMP matr	c 159	11.7	527	22	AAK76668	Human immune/haema
87	16	12.5	2623	25	AAZ20473	Matrix metalloprot	c 160	11.7	527	22	AAK76669	Human immune/haema
88	16	12.5	3152	22	AAKS1665	Human polynucleoti	c 161	11.7	527	22	AAK76670	Human immune/haema
89	16	12.5	3375	22	AAH17675	Human cDNA sequenc	c 162	11.7	536	22	AAK59981	Human immune/haema
90	16	12.5	3381	24	ABA93720	Human differentiat	c 163	11.7	565	22	AAH10992	Human cDNA clone (
91	16	12.5	3385	25	ABX05007	Human novel polynu	c 164	11.7	588	22	AAH10992	Human foetal liver
92	16	12.5	3449	22	AAK66521	Human immune syste	c 165	11.7	588	22	ABA30473	Probe #8939 for ge
93	16	12.5	3470	24	ABT03024	Human breast speci	c 166	11.7	588	22	AAK11702	Human brain expres
94	16	12.5	3748	19	AAV20627	Sf9 alpha-mannosid	c 167	11.7	588	22	AAK11702	Human bone marrow
95	16	12.5	3748	19	AAV16776	Sf9 alpha-mannosid	c 168	11.7	588	22	AAK118257	Probe #8190 for ge
96	16	12.5	3991	23	ABL07442	Drosophila melanog	c 169	11.7	588	22	AAI43306	Probe #81992 used
97	16	12.5	4263	23	AAAS93470	DNA encoding novel	c 170	11.7	588	23	ABS37127	Human liver single
98	16	12.5	4440	21	AAAS92169	Human POLG gene nu	c 171	11.7	588	24	ABS11430	Human genome-deriv
99	16	12.5	4465	24	AAI38925	Human polymerase g	c 172	11.7	603	23	ABV52142	Human prostate exp
100	16	12.5	4467	21	AAAC76835	Human ORFX ORF2390	c 173	11.7	616	23	ABV52101	Human prostate exp
101	16	12.5	5256	22	AAAO9193	Human matrix metal	c 174	11.7	633	22	AAH66069	C glutamicum codin
102	16	12.5	5256	22	AAK52649	Human polynucleoti	c 175	11.7	690	18	AAK30728	Streptococcus pneu
103	16	12.5	6160	24	ABL70233	Chemically treated	c 176	11.7	690	19	AAK98576	DNA encoding S. pn
104	16	12.5	6160	24	ABK31272	Signal transductio	c 177	11.7	756	22	AAH71525	Human cervical can
105	16	12.5	11893	23	ABM30327	Drosophila melanog	c 178	11.7	766	22	AAK91332	Human polynucleoti
106	16	12.5	13852	22	ABAI17744	Human nervous syst	c 179	11.7	910	21	AAK77380	Human ORFX ORF2935
107	16	12.5	13852	22	ABAI17745	Human nervous syst	c 180	11.7	931	22	ABAI4518	Human nervous syst
108	16	12.5	13852	22	ABAI1395	Human nervous syst	c 181	11.7	931	22	ABAI18941	Human nervous syst
109	16	12.5	13852	22	ABAI1396	Human nervous syst	c 182	11.7	1002	17	AAK13163	SCSV segment 4. S
110	16	12.5	15648	23	ABL30326	Drosophila melanog	c 183	11.7	1089	24	ABZ11808	Human polynucleoti
111	16	12.5	32186	22	AAK34422	Human DNA for a no	c 184	11.7	1089	24	ABZ11808	C glutamicum codin
112	16	12.5	75384	22	AAK55590	Human immune/haema	c 185	11.7	1143	22	ABAO8219	Human ovarian and
113	16	12.5	100301	24	ABQ88176	Human osteoblast d	c 186	11.7	1143	22	AAI07527	Human reproductiv
114	15	11.7	148	25	ABX54598	Bovine EST associ	c 187	11.7	1188	23	AAK89552	DNA encoding novel
115	15	11.7	193	22	ABAG7549	Human foetal liver	c 188	11.7	1189	23	AAK87858	DNA encoding novel
116	15	11.7	193	22	AAK41718	Human bone marrow	c 189	11.7	1213	21	AAK76807	Human ORFX ORF2362
117	15	11.7	193	22	AAI122475	Probe #12408 for g	c 190	11.7	1239	23	AAK85195	DNA encoding novel
118	15	11.7	193	22	AAI17766	Probe #16452 used	c 191	11.7	1241	22	AAH18491	Human cDNA sequenc
119	15	11.7	193	22	AAI08159	Probe #8150 used t	c 192	11.7	1248	22	AAK71603	Corynebacterium gl
120	15	11.7	193	22	ABSA1310	Human liver single	c 193	11.7	1248	22	AAK71603	Corynebacterium gl
121	15	11.7	193	24	ABSI1727	Human genome-deriv	c 194	11.7	1254	23	AAK89553	DNA encoding novel
122	15	11.7	212	21	AAK32125	Human secreted pro	c 195	11.7	1264	21	AAK73898	Human secreted pro
123	15	11.7	241	21	AAK13384	Human secreted pro	c 196	11.7	1289	22	AAK44602	Human full-length
124	15	11.7	244	25	ABX27048	Human ORF107 codin	c 197	11.7	1289	21	AAK93512	Human secreted pro
125	15	11.7	245	24	ABQ98300	Human ORF107 codin	c 198	11.7	1299	23	AAK93512	DNA encoding novel
126	15	11.7	263	21	AAK04793	Human secreted pro	c 199	11.7	1303	24	ABQ69050	Listeria monocytog
127	15	11.7	263	21	AAK04902	Human secreted pro	c 200	11.7	1413	20	AAK75834	H. pylori outer me
128	15	11.7	300	25	ABX06912	S. pneumoniae type	c 201	11.7	1563	21	AAK49681	Arabidopsis thalia
129	15	11.7	303	25	ABX06308	S. pneumoniae type	c 202	11.7	1584	24	ABZ32511	Candida albicans e
130	15	11.7	305	24	ABN23277	Human ORFX polynuc	c 203	11.7	1584	24	ABZ32511	Candida albicans e
131	15	11.7	307	23	ABV20246	Human prostate exp	c 204	11.7	1617	23	AAK76528	Haemophilus influe
132	15	11.7	307	23	ABV26077	Human prostate exp	c 205	11.7	1698	22	AAK53391	Haemophilus influe
133	15	11.7	312	23	ABV30684	Human prostate exp	c 206	11.7	1698	23	AAK53391	Human cervical can
134	15	11.7	312	23	ABV39652	Human prostate exp	c 207	11.7	1698	23	ABV24391	Human prostate exp
135	15	11.7	320	25	ABZ71684	Breast specific nu	c 208	11.7	1764	22	AAH15352	Human cDNA sequenc
136	15	11.7	323	23	ABV09508	Human prostate exp	c 209	11.7	1866	22	AAH15352	Human secreted pro
137	15	11.7	341	24	ABK09499	Human prostate exp	c 210	11.7	1886	25	ACC50434	Human secreted pro
138	15	11.7	345	24	ABK76476	Bacillus lichenifo	c 211	11.7	1902	21	AAZ46411	Pleurotus cornuocp
139	15	11.7	353	25	ABX21497	Human GDP-mannose	c 212	11.7	1951	22	AAI57898	Human polynucleoti
140	15	11.7	359	22	AAI80492	Human polynucleoti	c 213	11.7	1952	22	AAI59684	Human polynucleoti
141	15	11.7	395	24	ABL67041	Thyroid cancer rel	c 214	11.7	1957	22	AAK90585	CNA encoding huma
142	15	11.7	397	22	AAK61902	Human immune/haema	c 215	11.7	1957	24	AAK36755	Human RING finger
143	15	11.7	410	22	AAI87862	Human polynucleoti	c 216	11.7	1966	24	ABI99599	Mouse ischaemic co
144	15	11.7	423	22	AAK78273	Human immune/haema	c 217	11.7	1986	16	AAQ97797	Clone ptdelta2 cod
145	15	11.7	447	22	AAI14545	Probe #4478 for ge	c 218	11.7	2084	23	ABL10239	Drosophila melanog
146	15	11.7	447	22	AAI04363	Probe #4354 used t	c 219	11.7	2100	24	ABZ32348	Candida albicans e
147	15	11.7	458	25	ABZ71685	Breast specific nu	c 220	11.7	2199	22	AAH17620	Human cDNA sequenc
148	15	11.7	467	21	AAK40981	Zea mays DNA fragm	c 221	11.7	2283	21	AAK52791	Candida albicans e
149	15	11.7	479	22	ABAS4937	Human foetal liver	c 222	11.7	2283	21	AAK48526	Arabidopsis thalia
150	15	11.7	479	22	AAK28663	Human bone marrow	c 223	11.7	2376	24	ABN79825	Fungal ZBC gene se
151	15	11.7	479	22	AAI13240	Probe #1173 for ge	c 224	11.7	2478	22	AAI60057	Human polynucleoti
152	15	11.7	479	22	AAI34592	Probe #3278 used t	c 225	11.7	2528	25	ABZ35360	Human secretory po
153	15	11.7	479	22	AAI03143	Probe #3134 used t	c 226	11.7	2539	23	AAK82625	DNA encoding novel
154	15	11.7	479	23	AAK28262	Human liver single	c 227	11.7	2609	22	AAK91895	Human secreted pro

228	15	11.7	2622	22	AA158271	Human polynucleoti	301	15	11.7	23030	22	AA106475	Human reproductive
c 229	15	11.7	2633	24	AB513155	cDNA encoding huma	c 302	15	11.7	23030	22	AAK64779	Human immune/haema
c 230	15	11.7	2639	23	ABL03902	Drosophila melanog	303	15	11.7	23030	22	ABX59056	cDNA encoding nove
c 231	15	11.7	2720	23	AB110489	Drosophila melanog	304	15	11.7	23670	24	ABA97077	Tomato Hero gene c
c 232	15	11.7	2783	20	AAZ06969	Murine gamma-amino	305	15	11.7	30420	22	ABL136067	Human musculoskele
c 233	15	11.7	2883	22	AC931924	Murine GABA-B-Ria	306	15	11.7	30420	22	AA540588	DNA encoding human
c 234	15	11.7	2962	23	ABX63532	Human cDNA #532 di	307	15	11.7	30420	22	AA106475	Human reproductive
c 235	15	11.7	2964	23	AB247432	Drosophila melanog	308	15	11.7	30420	22	ABX59056	cDNA encoding nove
c 236	15	11.7	3001	21	AAH51755	Chromosome 13q31-q	309	15	11.7	31730	22	ABL17445	Human musculoskele
c 237	15	11.7	3023	18	AAAT67241	Pasteurella haemol	310	15	11.7	31730	25	ABX60433	cDNA encoding nove
c 238	15	11.7	3168	24	ABQ91991	Human NF-kB activa	c 311	15	11.7	33923	22	AAK67071	Human immune/haema
c 239	15	11.7	3168	24	ABQ91992	Human NF-kB activa	c 312	15	11.7	4861	24	AA520000	DNA encoding pyrid
c 240	15	11.7	3470	24	ABQ11026	Listeria monocytog	c 313	15	11.7	90220	24	ABR83576	Human cDNA differe
c 241	15	11.7	3810	23	AA599949	DNA encoding novel	c 314	15	11.7	125910	21	AAAC64370	Human KCM05 (KCM6q
c 242	15	11.7	3810	23	AA592574	DNA encoding novel	c 315	15	11.7	178896	24	ABQ88146	Human osteoblast d
c 243	15	11.7	3885	23	AA585580	DNA encoding novel	c 316	15	11.7	203654	25	ABX16034	Human gene encodin
c 244	15	11.7	3895	23	AA568999	DNA encoding novel	c 317	15	11.7	207433	23	ABZ72040	Gene 216 H194BAC10
c 245	15	11.7	3970	23	ABL15315	Drosophila melanog	c 318	15	11.7	207433	23	ABZ72040	BAC109L22 DNA seq
c 246	15	11.7	4058	23	ABL03757	Drosophila melanog	c 319	15	11.7	341511	24	ABX74891	Genomic DNA encodi
c 247	15	11.7	4205	23	AA580087	DNA encoding novel	c 320	15	11.7	349980	22	AAH41223	Pyrococcus abyssi
c 248	15	11.7	4205	23	AA585575	DNA encoding novel	c 321	15	11.7	349980	22	AAH68527	C glutamicum codin
c 249	15	11.7	4205	23	AA588138	DNA encoding novel	c 322	15	11.7	506286	24	ABX55320	Human transporter
c 250	15	11.7	4290	23	AA588152	DNA encoding novel	c 323	15	11.7	567571	25	AD53224	Human chromosome 3
c 251	15	11.7	4309	24	ABQ99288	Human coding seque	c 324	15	11.7	1082138	21	AA22305	Arabidopsis thalia
c 252	15	11.7	4365	24	ABX33497	Murine GABA-B rece	c 325	15	11.7	1830121	17	AA42063	Haemophilus influ
c 253	15	11.7	4365	25	ABZ68445	Nucleotide sequenc	c 326	15	11.7	2162598	25	ABX56454	Streptococcus pneu
c 254	15	11.7	4387	24	ABN59837	Novel human coding	c 327	15	11.7	2162598	25	ABX56454	Streptococcus pneu
c 255	15	11.7	5086	22	AA070355	Human KIAA0211 cDN	c 328	15	11.7	2944528	24	ABA03041	Listeria monocytog
c 256	15	11.7	5086	24	ABK83718	Human cDNA differe	c 329	14	10.9	24	24	AAH47936	Human sigma-54 fac
c 257	15	11.7	5813	22	AAH57490	Human liver cell s	c 330	14	10.9	25	19	AAZ09536	Human biallelic po
c 258	15	11.7	5926	24	ABL33653	Human immune syste	c 331	14	10.9	27	25	ABZ69758	M. jannaschii Ala5
c 259	15	11.7	6049	24	ABL32228	Human immune syste	c 332	14	10.9	60	24	ABN48264	Human spliced tran
c 260	15	11.7	6052	22	AA59391	Mouse cDNA encodin	c 333	14	10.9	69	25	ABZ69757	M. jannaschii Ala5
c 261	15	11.7	6075	22	AA59389	Mouse cDNA encodin	c 334	14	10.9	76	25	ABX26930	Human GDP-mannose
c 262	15	11.7	6120	22	AA59390	Mouse cDNA encodin	c 335	14	10.9	82	22	AAK47757	Human bone marrow
c 263	15	11.7	6191	24	ABN80160	Human chemically m	c 336	14	10.9	82	22	AAI53590	Probe #22276 used
c 264	15	11.7	6191	24	ABL70281	Chemically treated	c 337	14	10.9	82	23	ABX47486	Human liver single
c 265	15	11.7	6191	24	ABL33216	Human immune syste	c 338	14	10.9	114	19	AAH11882	Human biallelic po
c 266	15	11.7	6191	24	ABK13106	Signal transductio	c 339	14	10.9	170	24	ABL77410	Human ovarian canc
c 267	15	11.7	6217	24	ABL32938	Human immune syste	c 340	14	10.9	171	19	AAH12784	Human biallelic po
c 268	15	11.7	6734	22	AA104855	Human reproductive	c 341	14	10.9	175	22	ABX49596	Human breast cell
c 269	15	11.7	6734	23	ABL97749	Human testicular a	c 342	14	10.9	175	22	ABA67504	Human foetal liver
c 270	15	11.7	6746	23	ABL03756	Drosophila melanog	c 343	14	10.9	175	22	ABA67504	Probe #13049 for g
c 271	15	11.7	6750	23	ABL15314	Drosophila melanog	c 344	14	10.9	175	22	ABX45483	Human brain expres
c 272	15	11.7	6756	22	AA104856	Human reproductive	c 345	14	10.9	175	22	AAK15931	Human bone marrow
c 273	15	11.7	6756	23	ABL97750	Human testicular a	c 346	14	10.9	175	22	AAK1673	Human bone marrow
c 274	15	11.7	6999	23	AA585200	DNA encoding novel	c 347	14	10.9	175	22	AAI22427	Probe #12360 for g
c 275	15	11.7	7100	20	AAK24980	Pacillus subtilis	c 348	14	10.9	175	22	AAI47718	Probe #16404 used
c 276	15	11.7	7287	22	AAH57428	Human intestine ce	c 349	14	10.9	175	23	AAI08113	Probe #8104 used t
c 277	15	11.7	7481	18	AAV74547	Staphylococcus aur	c 350	14	10.9	175	23	ABX41262	Human liver single
c 278	15	11.7	7721	22	AAK65240	Human immune/haema	c 351	14	10.9	178	24	ABX15682	Human genome-deriv
c 279	15	11.7	7923	22	AAK65237	Human immune/haema	c 352	14	10.9	178	21	AAK21311	Human secreted pro
c 280	15	11.7	10011	19	AAV52209	Streptococcus pneu	c 353	14	10.9	209	21	AAK44735	Human secreted exp
c 281	15	11.7	10036	23	ABL10498	Drosophila melanog	c 354	14	10.9	231	21	AAK09239	Human secreted pro
c 282	15	11.7	11458	22	AAZ0414	P. syringae pv.tom	c 355	14	10.9	235	21	AAK28181	Human secreted pro
c 283	15	11.7	11467	23	ABL15274	Drosophila melanog	c 356	14	10.9	255	22	AAH82136	Rat differential t
c 284	15	11.7	12044	23	ABL10238	Drosophila melanog	c 357	14	10.9	259	24	ABN26999	Human ORFX polynuc
c 285	15	11.7	12044	23	ABL10238	Drosophila melanog	c 358	14	10.9	263	25	ABX84830	Corn ear-derived p
c 286	15	11.7	12174	22	AAK90854	Human digestive sy	c 359	14	10.9	267	25	ABX83421	Corn ear-derived p
c 287	15	11.7	12280	22	AB119083	Human nervous syst	c 360	14	10.9	298	25	ACA56999	Human adipocyte Se
c 288	15	11.7	12804	22	AAK76794	Human immune/haema	c 361	14	10.9	300	21	AAK01030	Human colon cancer
c 289	15	11.7	12988	22	AAK65240	Human nervous syst	c 362	14	10.9	310	25	ABX75378	Mouse cytokine mod
c 290	15	11.7	12988	25	AAK74054	Secreted protein g	c 363	14	10.9	314	22	AAK56483	Human immune/haema
c 291	15	11.7	13727	22	AAK31857	Human immune/haema	c 364	14	10.9	323	24	ABL77478	Human ovarian canc
c 292	15	11.7	14894	22	AAK81564	Human immune/haema	c 365	14	10.9	338	20	AAV87492	EST clone BU65. H
c 293	15	11.7	15363	19	AAV52272	Streptococcus pneu	c 366	14	10.9	341	19	ABX53318	DNA encoding a Sta
c 294	15	11.7	15525	22	AAK81856	Human immune/haema	c 367	14	10.9	349	25	ABX65200	Human gene trapped
c 295	15	11.7	17993	24	AAZ46721	Human immune/haema	c 368	14	10.9	353	22	AAI05149	Human reproductive
c 296	15	11.7	21477	22	AAK66626	Human transporter	c 369	14	10.9	353	22	AAI87467	Human polynucleoti
c 297	15	11.7	21480	22	AAK66625	Human immune/haema	c 370	14	10.9	355	23	ABL98036	Human testicular a
c 298	15	11.7	22636	22	AAK31855	Human immune/haema	c 371	14	10.9	356	22	ABX19762	Human GDP-mannose
c 299	15	11.7	23030	22	AAU36068	Human musculoskele	c 372	14	10.9	356	22	ABX34050	Human cDNA encodin
c 300	15	11.7	23030	22	AA540587	DNA encoding human	c 373	14	10.9	358	22	AAI86251	Bovine EST associa

374	14	10.9	363	22	AAK5528	Human immune/haema	C 447	14	10.9	471	25	ABZ20348	Group III CDNA can
375	14	10.9	365	22	AAK56311	Human immune/haema	C 448	14	10.9	472	22	AAH13403	Human CDNA clone (
376	14	10.9	368	22	AAK59069	Human cancer relat	C 449	14	10.9	482	21	AAK89643	Exo57 nucleotide s
377	14	10.9	375	16	AAK722131	Human gene signatu	C 450	14	10.9	482	24	AAK63353	Rat sequence diffe
378	14	10.9	375	20	AAK99877	Monoclonal antibod	C 451	14	10.9	491	22	AAK05082	Human brain expres
379	14	10.9	378	22	AAK67773	Human immune/haema	C 452	14	10.9	495	22	AAH12417	Human CDNA clone (
380	14	10.9	378	23	ABV45949	Human prostate exp	C 453	14	10.9	503	21	AAK09306	Human secreted pro
381	14	10.9	380	24	ABL33385	Human ovarian can	C 454	14	10.9	509	22	AAI86316	Human polynucleoti
382	14	10.9	382	25	ABX21678	Human GDP-mannose	C 455	14	10.9	511	21	AAK52960	Arabidopsis thalia
383	14	10.9	383	24	ABL62348	Colon adenocarcino	C 456	14	10.9	513	24	ABN60212	Human cancer relat
384	14	10.9	383	24	ABL65156	Lung cancer relate	C 457	14	10.9	515	24	ABN60450	Human cancer relat
385	14	10.9	383	24	ABL66239	Lung cancer relate	C 458	14	10.9	519	23	AAK92310	DNA encoding novel
386	14	10.9	383	24	ABL66834	Lung cancer relate	C 459	14	10.9	520	22	AAK61002	Human immune/haema
387	14	10.9	383	24	ABL67495	Thyroid cancer rel	C 460	14	10.9	526	23	ABV53261	Human prostate exp
388	14	10.9	387	22	ABA72754	Human foetal liver	C 461	14	10.9	531	21	AAK43446	Arabidopsis thalia
389	14	10.9	387	22	AAK21184	Human brain expres	C 462	14	10.9	531	24	ABZ12452	Arabidopsis thalia
390	14	10.9	387	22	AAK47340	Human bone marrow	C 463	14	10.9	531	24	ABK45436	cDNA encoding colo
391	14	10.9	387	22	AAK53177	Probe #21863 used	C 464	14	10.9	532	23	ABV60948	Human prostate exp
392	14	10.9	387	22	ABK57085	Human liver single	C 465	14	10.9	542	22	AAH11386	Human CDNA clone (
393	14	10.9	389	22	AAK53787	Murine translation	C 466	14	10.9	548	22	AAK35202	Human musculoskele
394	14	10.9	389	24	ABN78958	Human ORF3905 cDNA	C 467	14	10.9	548	25	ABX58190	cDNA encoding nove
395	14	10.9	390	22	AAK67771	Human immune/haema	C 468	14	10.9	549	22	ABK60215	Human foetal liver
396	14	10.9	390	22	AAK67772	Human immune/haema	C 469	14	10.9	549	22	AAK08492	Human brain expres
397	14	10.9	391	24	ABL30335	Human ovarian can	C 470	14	10.9	549	22	AAK34374	Human bone marrow
398	14	10.9	398	22	AAK20051	Human breast cance	C 471	14	10.9	549	22	AAI40098	Probe #8784 used t
399	14	10.9	398	22	AAK26400	Human breast cance	C 472	14	10.9	549	23	ABK34157	Human liver single
400	14	10.9	400	22	AAK06120	Human reproductive	C 473	14	10.9	556	22	ABD10130	Mouse channel indu
401	14	10.9	400	23	ABL98685	Human testicular a	C 474	14	10.9	556	24	ABK35064	Murine CDNA isolat
402	14	10.9	401	24	ABT07595	Human breast cance	C 475	14	10.9	567	24	ABN67071	Streptococcus poly
403	14	10.9	401	25	ABK41520	Bovine EST associa	C 476	14	10.9	575	22	ABA26757	Human nervous syst
404	14	10.9	405	25	ACA57207	Human adipocyte Se	C 477	14	10.9	578	21	AAK09032	Fusarium venenatum
405	14	10.9	407	21	AAK04137	Human secreted pro	C 478	14	10.9	582	22	AAK64015	Human immune/haema
406	14	10.9	407	24	ABK54059	Human head and nec	C 479	14	10.9	589	23	ABV53769	Human prostate exp
407	14	10.9	408	25	ABK38446	Bovine EST associa	C 480	14	10.9	593	22	ABK34792	Human nervous syst
408	14	10.9	411	23	ABV17035	Human prostate exp	C 481	14	10.9	596	22	AAK34792	Human bone marrow
409	14	10.9	411	24	ABN65112	Human cancer relat	C 482	14	10.9	596	22	AAI40512	Probe #9198 used t
410	14	10.9	413	22	AAK16934	Human breast cance	C 483	14	10.9	596	23	ABK34559	Human liver single
411	14	10.9	414	22	AAH36049	Human colon cancer	C 484	14	10.9	598	23	ABV53519	Human prostate exp
412	14	10.9	414	22	AAH362528	Human ovarian can	C 485	14	10.9	606	22	ABD05524	Human secreted pro
413	14	10.9	415	22	AAH17729	Human breast cance	C 486	14	10.9	608	22	AAK74724	Human immune/haema
414	14	10.9	419	22	AAH10471	Human breast cance	C 487	14	10.9	608	22	AAK86387	Human immune/haema
415	14	10.9	420	21	AAK30605	Human colon cancer	C 488	14	10.9	612	23	ABV50923	Human prostate exp
416	14	10.9	421	24	ABN37166	Gene #3664 used to	C 489	14	10.9	612	24	ABK61427	Prostate specific
417	14	10.9	427	22	ABK44898	Human breast cell	C 490	14	10.9	614	24	ABK55340	Human colon cancer
418	14	10.9	427	22	ABK55364	Human foetal liver	C 491	14	10.9	621	24	ABK66576	Streptococcus poly
419	14	10.9	427	22	ABK35082	Probe #3558 for ge	C 492	14	10.9	628	21	AAK69419	Human secreted pro
420	14	10.9	427	22	AAK3605	Human brain expres	C 493	14	10.9	630	21	AAK53502	Arabidopsis thalia
421	14	10.9	427	22	AAK29066	Human bone marrow	C 494	14	10.9	630	22	AAK61184	Human immune/haema
422	14	10.9	427	22	AAK13654	Human prostate exp	C 495	14	10.9	648	24	ABN62696	Human cancer relat
423	14	10.9	427	22	AAI35011	Probe #3587 for ge	C 496	14	10.9	668	21	AAI2700	Aspergillus oryzae
424	14	10.9	427	22	AAI03537	Probe #3697 used t	C 497	14	10.9	674	22	AAI13004	Human breast cance
425	14	10.9	427	23	ABK28683	Probe #3528 used t	C 498	14	10.9	674	22	AAK57749	Human immune/haema
426	14	10.9	427	24	ABK36612	Human liver single	C 499	14	10.9	679	22	AAK71056	Human immune/haema
427	14	10.9	430	25	ABK53432	Human genome-deriv	C 500	14	10.9	682	21	AAA02625	Human colon cancer
428	14	10.9	431	22	AAK10654	Bovine EST associa	C 501	14	10.9	690	22	AAH07330	Human CDNA clone (
429	14	10.9	432	22	ABK44445	Human breast cance	C 502	14	10.9	693	23	ABL30361	Drosophila melanog
430	14	10.9	432	22	ABK54892	Human breast cell	C 503	14	10.9	719	21	AAK09492	Human secreted pro
431	14	10.9	432	22	ABK24658	Human foetal liver	C 504	14	10.9	722	22	AAK52376	S. epidermidis ope
432	14	10.9	432	22	AAK03188	Probe #3124 for ge	C 505	14	10.9	744	22	AAK59397	Mouse potassium-ch
433	14	10.9	432	22	AAK28618	Human brain expres	C 506	14	10.9	750	15	AAK55231	Grapevine ribosoma
434	14	10.9	432	22	AAK13192	Human bone marrow	C 507	14	10.9	775	22	AAI65823	Nucleotide sequenc
435	14	10.9	432	22	AAI34544	Probe #3125 for ge	C 508	14	10.9	775	22	AAI21874	Human breast cance
436	14	10.9	432	22	AAI03097	Probe #3230 used t	C 509	14	10.9	785	22	AAH05024	Human CDNA clone (
437	14	10.9	432	22	ABK28214	Probe #3088 used t	C 510	14	10.9	792	22	AAK29157	CDNA encoding for
438	14	10.9	432	24	ABK30129	Human liver single	C 511	14	10.9	792	24	ABK68297	Human prostate exp
439	14	10.9	433	25	ABK52986	Human genome-deriv	C 512	14	10.9	792	24	ABK08957	Human prostate exp
440	14	10.9	434	21	AAK06496	Bovine EST associa	C 513	14	10.9	808	24	ABT08071	Human breast speci
441	14	10.9	434	23	AAK72286	Human secreted pro	C 514	14	10.9	819	19	AAV01417	Mouse T-cell recep
442	14	10.9	440	21	AAK31408	DNA encoding novel	C 515	14	10.9	821	24	ABQ38178	Oligonucleotide fo
443	14	10.9	453	24	ABV97214	Plant microstelli	C 516	14	10.9	821	24	ABQ38179	Oligonucleotide fo
444	14	10.9	461	23	ABK46830	Human pancreatic c	C 517	14	10.9	835	22	AAH00997	Enterococcus raffi
445	14	10.9	462	22	AAK56204	Human prostate exp	C 518	14	10.9	852	24	ABN91338	Staphylococcus epi
446	14	10.9	467	22	AAK35563	Human immune/haema	C 519	14	10.9	856	24	ABK61428	Prostate specific

520	14	10.9	857	24	AB52722	Dihydrofolate redu	c 593	14	10.9	1252	21	AAF22419	Human secreted pro
c 521	14	10.9	858	24	ABN70715	Streptococcus poly	c 594	14	10.9	1252	25	ABT16866	Human secreted pro
522	14	10.9	862	24	AB089299	Human prostate exp	c 595	14	10.9	1252	25	ABZ67292	Human secreted pro
523	14	10.9	873	23	AB577431	DNA encoding novel	c 596	14	10.9	1263	24	ABQ90455	M. capaulatus gene
c 524	14	10.9	892	24	AB666787	Helicobacter pylor	c 597	14	10.9	1279	24	ABC39313	Arabidopsis thalia
c 525	14	10.9	917	25	AB269768	Orthogonal aminoac	c 598	14	10.9	1288	24	ABZ16136	Arabidopsis thalia
c 526	14	10.9	917	25	AB269769	Orthogonal aminoac	c 599	14	10.9	1302	24	AB555941	DNA topoisomerase
c 527	14	10.9	917	25	AB269770	Orthogonal aminoac	600	14	10.9	1324	25	ABQ82966	Human lung specifi
c 528	14	10.9	917	25	AB269771	Orthogonal aminoac	601	14	10.9	1346	21	AA081757	Human secreted pro
c 529	14	10.9	917	25	AB269772	Orthogonal aminoac	602	14	10.9	1346	25	ABZ73658	Secreted protein-e
c 530	14	10.9	917	25	AB269773	Orthogonal aminoac	603	14	10.9	1346	25	ABT16850	Human secreted pro
c 531	14	10.9	917	25	AB269774	Orthogonal aminoac	604	14	10.9	1346	25	ABZ67251	Human secreted pro
c 532	14	10.9	917	25	AB269775	Orthogonal aminoac	c 605	14	10.9	1347	22	AA523419	Candida albicans e
c 533	14	10.9	917	25	AB269776	Orthogonal aminoac	c 606	14	10.9	1347	24	ABZ31751	Candida albicans e
c 534	14	10.9	917	25	AB269777	Orthogonal aminoac	c 607	14	10.9	1354	22	AA526907	Human cDNA encodin
c 535	14	10.9	917	25	AB269778	Orthogonal aminoac	c 608	14	10.9	1354	24	ABO54643	Human ovarian anti
c 536	14	10.9	917	25	AB269779	Orthogonal aminoac	609	14	10.9	1367	21	AA336057	Arabidopsis thalia
c 537	14	10.9	917	25	AB269791	Orthogonal aminoac	610	14	10.9	1371	21	AA534995	Arabidopsis thalia
c 538	14	10.9	917	25	AB269792	Orthogonal aminoac	c 611	14	10.9	1371	22	AA526834	Human cDNA encodin
c 539	14	10.9	917	25	AB269793	Orthogonal aminoac	c 612	14	10.9	1376	22	AAF27230	Rat adrenomedullin
c 540	14	10.9	917	25	AB269794	Orthogonal aminoac	c 613	14	10.9	1376	22	AAF29140	Rat adrenomedullin
c 541	14	10.9	917	25	AB269795	Orthogonal aminoac	c 614	14	10.9	1398	24	ADD28486	Bacillus subtilis
c 542	14	10.9	921	25	AB269767	Orthogonal aminoac	615	14	10.9	1406	24	ABQ77917	Human laminin rece
c 543	14	10.9	921	25	AB269780	Orthogonal aminoac	c 616	14	10.9	1414	22	AAK79529	Human immune/haema
c 544	14	10.9	921	25	AB269781	Orthogonal aminoac	617	14	10.9	1414	24	ABL68937	Kidney cancer rela
c 545	14	10.9	921	25	AB269782	Orthogonal aminoac	c 618	14	10.9	1436	21	AAA40759	Rat strain Brown N
c 546	14	10.9	921	25	AB269783	Orthogonal aminoac	c 619	14	10.9	1445	21	AAA40758	Rat strain Wistar
c 547	14	10.9	921	25	AB269784	Orthogonal aminoac	c 620	14	10.9	1469	24	ABK62592	Rat sequence diffe
c 548	14	10.9	921	25	AB269785	Orthogonal aminoac	621	14	10.9	1483	22	AAK68118	Human immune/haema
c 549	14	10.9	921	25	AB269786	Orthogonal aminoac	622	14	10.9	1510	20	AAK20426	Human secreted pro
c 550	14	10.9	921	25	AB269787	Orthogonal aminoac	623	14	10.9	1557	20	AAK61524	B. burgdorferi ant
c 551	14	10.9	921	25	AB269788	Orthogonal aminoac	c 624	14	10.9	1561	22	AA523809	Candida albicans D
c 552	14	10.9	921	25	AB269789	Orthogonal aminoac	c 625	14	10.9	1569	23	ABL04647	Drosophila melanog
c 553	14	10.9	921	25	AB269790	Orthogonal aminoac	c 626	14	10.9	1575	23	AA583383	Ornithine decarboxyl
554	14	10.9	936	21	AAZ50162	Soybean Aclerylsu	c 627	14	10.9	1575	23	AA589427	DNA encoding novel
555	14	10.9	946	22	AAZ23280	Human breast cance	c 628	14	10.9	1575	23	AA593361	DNA encoding novel
556	14	10.9	947	24	ABZ17023	Arabidopsis thalia	c 629	14	10.9	1576	22	AAF24154	Human secreted pro
557	14	10.9	948	18	AAK31985	DNA encoding a Sta	630	14	10.9	1581	22	AA587358	Human developmen
558	14	10.9	965	22	AAK68117	Human immune/haema	c 631	14	10.9	1587	22	AA523395	Candida albicans e
c 559	14	10.9	990	24	ABL63703	Breast cancer rela	c 632	14	10.9	1587	24	ABZ31744	Candida albicans e
c 560	14	10.9	990	24	ABL64099	Breast cancer rela	633	14	10.9	1599	22	AA591430	Human CHRM5 coding
c 561	14	10.9	990	24	ABL67237	Thyroid cancer rel	634	14	10.9	1599	23	AB198000	Non-endogenous hum
c 562	14	10.9	1011	23	ABL07427	Drosophila melanog	635	14	10.9	1599	24	ABK52225	Human cholinergic
c 563	14	10.9	1024	24	ABK66888	Helicobacter pylor	636	14	10.9	1599	25	ABZ42702	Human muscarinic a
564	14	10.9	1031	16	AAQ99786	Plant SAR gene pDP	637	14	10.9	1604	24	AB564396	Human cysteine sul
565	14	10.9	1031	20	AAV62802	Tobacco SAR CHX in	638	14	10.9	1617	20	AAK61523	B. burgdorferi ant
566	14	10.9	1031	20	AAV61686	Arabidopsis protei	639	14	10.9	1633	22	AAK34849	Human colon cancer
567	14	10.9	1031	22	AAK70142	Human immune/haema	c 640	14	10.9	1638	25	ABX05954	S. pneumoniae type
568	14	10.9	1045	22	AAK97365	Adult rat hippocam	c 641	14	10.9	1670	22	ABD08210	Human secreted pro
569	14	10.9	1046	21	AAK04764	Arabidopsis thalia	c 642	14	10.9	1670	22	ABD08210	Human secreted pro
c 570	14	10.9	1055	23	AAK04271	Drosophila melanog	c 643	14	10.9	1684	22	ABD08191	Human secreted pro
c 571	14	10.9	1077	23	AA575275	DNA encoding novel	c 644	14	10.9	1693	25	ABD47447	Tobacco cytochrome
c 572	14	10.9	1091	24	AAQ68536	Listeria monocytog	c 645	14	10.9	1701	25	ABZ18395	Group III cDNA can
c 573	14	10.9	1107	18	AAK68032	H. pylori flagella	c 646	14	10.9	1709	21	AAA40757	Rat SHR coding seq
c 574	14	10.9	1108	20	AAZ31913	Human helicase, Re	c 647	14	10.9	1720	25	ABX05065	Human novel polyu
c 575	14	10.9	1121	23	AA566401	DNA encoding novel	c 648	14	10.9	1745	25	ABD47446	Tobacco cytochrome
576	14	10.9	1123	20	AA573340	Human secreted pro	c 649	14	10.9	1750	7	AA607093	Sequence encoding
577	14	10.9	1129	24	AAK73730	Arabidopsis thalia	c 650	14	10.9	1750	12	AAQ11682	Sequence encoding
578	14	10.9	1147	24	AAK48251	Ehrlichia ruminant	651	14	10.9	1770	22	AAH18600	Human cDNA sequenc
579	14	10.9	1149	23	AAK67394	DNA encoding novel	c 652	14	10.9	1773	24	ABN70655	Streptococcus poly
c 580	14	10.9	1151	22	AAH98218	Human EST-derived	c 653	14	10.9	1786	22	AAH00842	Streptococcus pyog
c 581	14	10.9	1158	21	AAK48362	Arabidopsis thalia	c 654	14	10.9	1825	24	ABK85751	DNA encoding nicot
582	14	10.9	1174	22	AAH47934	Human sigma-54 fac	655	14	10.9	1844	20	AA224588	Human lung tumor a
c 583	14	10.9	1194	24	ABZ12459	Arabidopsis thalia	656	14	10.9	1844	21	AAK65827	Human lung cancer-
c 584	14	10.9	1209	23	AA593366	DNA encoding novel	657	14	10.9	1844	24	ABQ92232	Human lung cancer
c 585	14	10.9	1212	21	AAK79785	Human secreted pro	658	14	10.9	1844	24	ABL49046	Human lung tumour
586	14	10.9	1218	21	AA54996	Arabidopsis thalia	c 659	14	10.9	1847	24	ABT07596	Human breast cance
587	14	10.9	1219	21	AAK33899	Arabidopsis thalia	660	14	10.9	1848	24	AAK41389	Human NZMS-2 cDNA
588	14	10.9	1221	21	AAK72274	Drosophila odorant	c 661	14	10.9	1849	21	AAK48781	Arabidopsis thalia
589	14	10.9	1237	22	AAK65988	Human protein asso	c 662	14	10.9	1851	21	AAK37637	Arabidopsis thalia
c 590	14	10.9	1251	21	AAK59177	Human secreted pro	c 663	14	10.9	1864	22	AAH99713	Human protein enco
c 591	14	10.9	1251	25	ABT16858	Human secreted pro	c 664	14	10.9	1920	23	AA575888	DNA encoding novel
c 592	14	10.9	1251	25	ABZ67267	Human secreted pro	c 665	14	10.9	1923	24	ABN67224	Streptococcus poly



666	14	10.9	1935	18	AAV74659	Staphylococcus aur	C 739	14	10.9	3111	25	ABZ23269	Nucleotide sequenc
C 667	14	10.9	1935	22	AAK67546	Human immune/haema	C 740	14	10.9	3111	25	AAZ49450	Human kinase and p
C 668	14	10.9	1980	22	AAK63783	Human foetal liver	741	14	10.9	3141	22	AAI65824	Nucleotide sequenc
C 669	14	10.9	1980	22	AAK38019	Human bone marrow	742	14	10.9	3204	22	AAH54576	S. epidermidis gen
C 670	14	10.9	1980	22	AAI43904	Probe #12550 used	743	14	10.9	3205	22	AAO6848	Human reproductive
C 671	14	10.9	1980	23	ABS37637	Human liver single	744	14	10.9	3239	23	AAAS76858	DNA encoding novel
C 672	14	10.9	1980	24	ABS12021	Human genome-deriv	745	14	10.9	3254	23	AAH130360	Drosophila melanog
673	14	10.9	2000	24	ABZ16010	Arabidopsis thalia	C 746	14	10.9	3274	20	AAI13539	Enterococcus faeca
674	14	10.9	2000	24	ABZ17127	Arabidopsis thalia	C 747	14	10.9	3274	24	ABS99334	Enterococcus faeca
C 675	14	10.9	2125	22	AAI64631	Mouse MITP-origin	C 748	14	10.9	3284	22	AAZ77779	Human INTERCEPT 34
C 676	14	10.9	2150	23	ABL15069	Drosophila melanog	749	14	10.9	3286	24	AAAS94963	Human DNA sequence
C 677	14	10.9	2151	22	ABA15210	Human nervous syst	750	14	10.9	3326	24	AAZ75055	TR16-long receptor
C 678	14	10.9	2156	22	ABA15209	Human nervous syst	751	14	10.9	3390	22	AAZ75054	Fragment of pig oe
C 679	14	10.9	2178	24	ABK13908	Arabidopsis thalia	752	14	10.9	3460	17	AAZ75054	TR16-short recepto
C 680	14	10.9	2178	24	ABK13908	Arabidopsis thalia	753	14	10.9	3460	17	AAZ75054	Fragment of pig oe
C 681	14	10.9	2203	24	AAI48889	Human Pftaire fami	754	14	10.9	3518	24	ABN86358	Drosophila melanog
C 682	14	10.9	2206	23	ABL27709	Drosophila melanog	755	14	10.9	3546	22	AAH17870	Novel human protei
C 683	14	10.9	2222	22	AAK94265	Human full-length	C 756	14	10.9	3556	22	AAZ75055	Human cDNA sequenc
684	14	10.9	2253	10	AAZ92071	Human muscarinic a	C 757	14	10.9	3569	22	AAH54555	TR16-long receptor
C 685	14	10.9	2253	24	ABS98565	Human acetyl choli	758	14	10.9	3609	22	AAH54555	Drosophila melanog
C 686	14	10.9	2253	24	ABN67496	Streptococcus poly	C 759	14	10.9	3622	22	AAK73534	S. epidermidis gen
C 687	14	10.9	2261	22	AAA91429	Human CHRM5 gene s	760	14	10.9	3636	23	ABL11340	Human immune/haema
C 688	14	10.9	2261	25	ACA56578	Human signalling p	C 761	14	10.9	3692	22	AAZ74295	Drosophila melanog
C 689	14	10.9	2306	22	AAH14406	Human cDNA sequenc	C 762	14	10.9	3707	20	AAZ31910	S cerevisiae Yll02
C 690	14	10.9	2309	25	ABZ6401	Human secretory po	C 763	14	10.9	3708	24	ABS64845	Human tumour suppr
C 691	14	10.9	2340	24	ABA95468	Human natural kill	C 764	14	10.9	3715	22	AAZ75054	Human RecQ5 alpha
C 692	14	10.9	2354	17	AAZ63933	Lactococcus lactis	C 765	14	10.9	3716	23	ABL18078	Drosophila melanog
C 693	14	10.9	2360	22	AAZ07734	Human secreted pro	C 766	14	10.9	3721	23	ABL17356	Drosophila melanog
C 694	14	10.9	2370	21	AAZ45027	Human immune syst	C 767	14	10.9	3729	22	AAH54303	S. epidermidis gen
C 695	14	10.9	2378	24	ABA33412	Human cDNA sequenc	768	14	10.9	3758	25	ABZ42800	Human C-C chemokin
C 696	14	10.9	2386	24	ABA33412	Human cDNA sequenc	C 769	14	10.9	3758	25	ABZ42800	Streptococcus pneu
C 697	14	10.9	2403	22	AAH17531	Human chlamydia der	C 770	14	10.9	3758	25	ABZ42800	Streptococcus pneu
C 698	14	10.9	2436	21	AAZ40756	Human cDNA sequenc	C 771	14	10.9	3758	25	ABZ42800	Streptococcus pneu
C 699	14	10.9	2507	21	AAZ40756	Wild type Sprague-	C 772	14	10.9	3850	24	ABZ42800	Mouse ischaemic co
700	14	10.9	2507	21	AAZ40756	DNA encoding a rec	C 773	14	10.9	3850	24	ABZ42800	Mouse nervous syst
C 701	14	10.9	2533	23	ABL10925	DNA encoding a rec	C 774	14	10.9	3850	24	ABZ42800	Human nervous syst
702	14	10.9	2560	24	ABQ1204	Drosophila melanog	C 775	14	10.9	4018	25	ABZ65882	Nucleotide sequenc
703	14	10.9	2568	22	AAH15184	Protein expressed	C 776	14	10.9	4018	25	ABZ65882	DNA encoding novel
704	14	10.9	2568	22	AAH15184	Human cDNA sequenc	C 777	14	10.9	4018	25	ABZ65882	Human prostate exp
C 705	14	10.9	2580	18	AAZ86009	Human serine prote	C 778	14	10.9	4233	21	AAZ45600	Drosophila melanog
C 706	14	10.9	2582	24	ABK74020	Mouse gamma meltri	C 779	14	10.9	4233	21	AAZ45600	cDNA sequence of h
707	14	10.9	2582	24	ABK74020	Bacillus lichenifo	C 780	14	10.9	4233	21	AAZ45600	Japanese firefly h
708	14	10.9	2582	24	ABK74020	Human laminin B2 c	C 781	14	10.9	4233	21	AAZ45600	Drosophila melanog
C 709	14	10.9	2598	22	AAI58191	Human polynucleoti	C 782	14	10.9	4233	21	AAZ45600	Drosophila melanog
C 710	14	10.9	2598	24	ABZ20647	Human flavoprotein	C 783	14	10.9	4233	21	AAZ45600	Ehrlichia ruminant
C 711	14	10.9	2670	22	AAI58659	Human polynucleoti	C 784	14	10.9	4233	21	AAZ45600	Ehrlichia ruminant
C 712	14	10.9	2700	24	ABQ70610	Listeria monocytog	C 785	14	10.9	4233	21	AAZ45600	Drosophila melanog
713	14	10.9	2717	20	AAI13034	Enterococcus faeca	C 786	14	10.9	4233	21	AAZ45600	Aspergillus fumiga
714	14	10.9	2717	20	AAI13034	Enterococcus faeca	C 787	14	10.9	4233	21	AAZ45600	Drosophila melanog
715	14	10.9	2760	22	AAZ07710	Human secreted pro	C 788	14	10.9	4233	21	AAZ45600	Human immune/haema
C 716	14	10.9	2770	21	AAZ26323	Human secreted pro	C 789	14	10.9	4233	21	AAZ45600	Human immune/haema
C 717	14	10.9	2787	24	ABV74470	UGT and cobalamin	C 790	14	10.9	4233	21	AAZ45600	Human immune/haema
C 718	14	10.9	2790	24	ABA02168	Laminin 19-encodin	C 791	14	10.9	4233	21	AAZ45600	Human immune/haema
C 719	14	10.9	2818	21	AAZ45599	cDNA sequence of h	C 792	14	10.9	4233	21	AAZ45600	Human immune/haema
C 720	14	10.9	2835	17	AAZ28772	Human calcium sens	C 793	14	10.9	4233	21	AAZ45600	Human immune/haema
C 721	14	10.9	2835	17	AAZ28772	Human calcium sens	C 794	14	10.9	4233	21	AAZ45600	Human immune/haema
C 722	14	10.9	2855	24	ABZ9994	Human calcium sens	C 795	14	10.9	4233	21	AAZ45600	Human immune/haema
C 723	14	10.9	2865	24	ABQ1047	Human calcium sens	C 796	14	10.9	4233	21	AAZ45600	Human immune/haema
C 724	14	10.9	2894	24	ABQ70754	Human calcium sens	C 797	14	10.9	4233	21	AAZ45600	Human immune/haema
C 725	14	10.9	2926	23	ABL24510	3 isoform protein	C 798	14	10.9	4233	21	AAZ45600	Human immune/haema
C 726	14	10.9	2931	22	AAK37690	Listeria monocytog	C 799	14	10.9	4233	21	AAZ45600	Human immune/haema
C 727	14	10.9	2935	22	ABA76052	Drosophila melanog	C 800	14	10.9	4233	21	AAZ45600	Human immune/haema
C 728	14	10.9	2935	22	AAK17801	Human foetal liver	C 801	14	10.9	4233	21	AAZ45600	Human immune/haema
C 729	14	10.9	2935	22	AAK30728	Human brain expres	C 802	14	10.9	4233	21	AAZ45600	Human immune/haema
C 730	14	10.9	2935	22	AAI56711	Human bone marrow	C 803	14	10.9	4233	21	AAZ45600	Human immune/haema
C 731	14	10.9	2935	23	ABS50320	Probe #25397 used	C 804	14	10.9	4233	21	AAZ45600	Human immune/haema
C 732	14	10.9	2935	24	ABS24200	Human liver single	C 805	14	10.9	4233	21	AAZ45600	Human immune/haema
C 733	14	10.9	2967	23	ABN70173	Human genome-deriv	C 806	14	10.9	4233	21	AAZ45600	Human immune/haema
C 734	14	10.9	2976	24	ABN70173	Drosophila melanog	C 807	14	10.9	4233	21	AAZ45600	Human immune/haema
C 735	14	10.9	3000	23	ABN19677	Streptococcus poly	C 808	14	10.9	4233	21	AAZ45600	Human immune/haema
C 736	14	10.9	3001	21	AAH51765	Drosophila melanog	C 809	14	10.9	4233	21	AAZ45600	Human immune/haema
C 737	14	10.9	3011	23	ABQ7426	Chromosome 13q11-q	C 810	14	10.9	4233	21	AAZ45600	Human immune/haema
C 738	14	10.9	3066	24	ABQ66165	Drosophila melanog	C 811	14	10.9	4233	21	AAZ45600	Human immune/haema
	14	10.9	3111	24	ABN86357	Novel human protei		14	10.9	4233	21	AAZ45600	Human immune/haema

C 812	14	10.9	5870	22	AA501146	Interferon induced	C 885	14	10.9	11147	25	ABZ10007	Haematopoietic cel
C 813	14	10.9	5881	24	AB199651	Mouse ischaemic co	C 886	14	10.9	11147	25	ABZ10153	Haematopoietic cel
C 814	14	10.9	5993	24	ABU70599	Chemically treated	C 887	14	10.9	11294	23	ABL04652	Drosophila melanog
C 815	14	10.9	5993	24	AA561326	Human gene regulat	C 888	14	10.9	11474	21	AA53720	LPS core biosynthe
C 816	14	10.9	5993	24	ABK3402	Signal transductio	C 889	14	10.9	11474	25	ABT13665	Campylobacter jeju
C 817	14	10.9	6000	19	AV04770	Yeast transcriptio	C 890	14	10.9	11556	24	ABK69902	Human secreted pro
C 818	14	10.9	6000	23	ABL08655	Drosophila melanog	C 891	14	10.9	11943	22	AAK90588	Human digestive sy
C 819	14	10.9	6052	22	AA559391	Mouse cDNA encodin	C 892	14	10.9	11953	22	AAK90589	Human digestive sy
C 820	14	10.9	6103	24	ABL33691	Human immune syste	C 893	14	10.9	11962	24	ABS54194	OCF-associated seq
C 821	14	10.9	6106	22	AA546429	Tumour suppressor	C 894	14	10.9	11962	24	ABK90043	OCF related DNA se
C 822	14	10.9	6106	24	ABK40031	Human chemically p	C 895	14	10.9	11967	24	ABS54184	Mouse osteoclast p
C 823	14	10.9	6106	24	ABL33472	Human immune syste	C 896	14	10.9	11967	24	ABK90034	m608p-Lexicon clon
C 824	14	10.9	6124	23	ABL210176	Drosophila melanog	C 897	14	10.9	11967	24	AAI72584	Partial mouse OCP
C 825	14	10.9	6162	24	ABK31330	Signal transductio	C 898	14	10.9	12142	24	ABL33673	Human immune syste
C 826	14	10.9	6219	24	ABL32867	Human immune syste	C 899	14	10.9	12971	23	ABL11010	Drosophila melanog
C 827	14	10.9	6219	24	AA563325	Chemically pretrea	C 900	14	10.9	13055	23	ABL25732	Drosophila melanog
C 828	14	10.9	6412	17	AAI28796	Human calcium sens	C 901	14	10.9	13165	22	ABA07140	Human pancreatic c
C 829	14	10.9	6412	19	AV06017	Human calcium prot	C 902	14	10.9	13165	22	AAK89348	Human digestive sy
C 830	14	10.9	6454	23	ABL11966	Drosophila melanog	C 903	14	10.9	13189	22	AAI06063	Human reproductive
C 831	14	10.9	6475	24	ABL34249	Human immune syste	C 904	14	10.9	13189	23	ABL98628	Human testicular a
C 832	14	10.9	6482	24	ABL22246	Chemically treated	C 905	14	10.9	13376	24	ABL32582	Human immune syste
C 833	14	10.9	6482	24	ABL22323	Chemically treated	C 906	14	10.9	13632	23	ABL07160	Drosophila melanog
C 834	14	10.9	6497	24	ABL32201	Human immune syste	C 907	14	10.9	13712	24	ABL33531	Human immune syste
C 835	14	10.9	6509	24	AA561086	Human gene regulat	C 908	14	10.9	14042	19	AAI28774	Human placenta cal
C 836	14	10.9	6509	24	ABK31189	Signal transductio	C 909	14	10.9	14042	19	AAV05995	Human placenta cal
C 837	14	10.9	6516	22	AA532715	Human genomic DNA	C 910	14	10.9	14044	17	AAI28776	Human parathyroid
C 838	14	10.9	6516	22	AAI07084	Human reproductive	C 911	14	10.9	14044	19	AAV05997	Human parathyroid
C 839	14	10.9	6580	23	ABL11888	Drosophila melanog	C 912	14	10.9	14080	17	AAI28775	Human kidney calci
C 840	14	10.9	6608	24	ABL34035	Human immune syste	C 913	14	10.9	14080	17	AAV05996	Human kidney calci
C 841	14	10.9	6638	22	AAK79592	Human immune/haema	C 914	14	10.9	14086	19	AAI28773	Human calcium sens
C 842	14	10.9	6638	22	AAK83770	Human immune/haema	C 915	14	10.9	14086	19	AAV05994	Human calcium prot
C 843	14	10.9	6874	22	AA545441	Chemically pretrea	C 916	14	10.9	14147	22	AA546743	Tumour suppressor
C 844	14	10.9	6874	24	ABN80231	Human chemically m	C 917	14	10.9	14147	24	ABK33955	Human DNA for stag
C 845	14	10.9	6874	24	ABL70382	Chemically treated	C 918	14	10.9	14216	22	AAK79131	Human immune/haema
C 846	14	10.9	6874	24	AA561334	Human gene regulat	C 919	14	10.9	14216	22	AAK79133	Human immune/haema
C 847	14	10.9	6874	24	ABK28290	DNA transcription	C 920	14	10.9	14244	22	AAH57510	Human kidney cell
C 848	14	10.9	6881	24	ABL33380	Human immune syste	C 921	14	10.9	14308	23	ABL11804	Drosophila melanog
C 849	14	10.9	6881	24	ABK28249	DNA transcription	C 922	14	10.9	14379	22	AAK79132	Human immune/haema
C 850	14	10.9	6925	24	ABL33388	Human immune syste	C 923	14	10.9	14392	23	AA568717	DNA encoding novel
C 851	14	10.9	7118	22	AA546382	Tumour suppressor	C 924	14	10.9	14392	25	ABZ34886	Coding sequence SE
C 852	14	10.9	7118	24	ABL32831	Human immune syste	C 925	14	10.9	14452	23	ABV22986	Human prostate exp
C 853	14	10.9	7168	24	ABU92231	Chemically treated	C 926	14	10.9	14455	23	ABV22821	Human prostate exp
C 854	14	10.9	7168	24	AAQ22318	Chemically treated	C 927	14	10.9	14456	23	ABV24304	Human prostate exp
C 855	14	10.9	7368	11	AAQ04013	Acetyl-CoA-carboxy	C 928	14	10.9	14460	24	ABA01440	Streptococcus ther
C 856	14	10.9	7420	24	ABK62595	Rat sequence diffe	C 929	14	10.9	14615	22	AA545704	Tumour suppressor
C 857	14	10.9	7461	22	AA530639	DNA encoding novel	C 930	14	10.9	14617	22	ABA15132	Human nervous syst
C 858	14	10.9	7461	22	AA528701	Genomic sequence #	C 931	14	10.9	14627	22	ABA19130	Human nervous syst
C 859	14	10.9	7461	25	ACA03402	DNA encoding human	C 932	14	10.9	15071	22	AAH53785	S. epidermidis ope
C 860	14	10.9	7528	20	AAI12992	Enterococcus faeca	C 933	14	10.9	15129	22	AAI03207	Human reproductive
C 861	14	10.9	7528	24	ABN98787	Enterococcus faeca	C 934	14	10.9	15255	22	AAK83136	Human immune/haema
C 862	14	10.9	7534	24	ABN80155	Human chemically m	C 935	14	10.9	15256	22	AAK83134	Human immune/haema
C 863	14	10.9	7733	22	AAH02136	Clostridium diffic	C 936	14	10.9	15838	22	ABK20613	Human nervous syst
C 864	14	10.9	7990	24	ABU54307	Chemically treated	C 937	14	10.9	16127	24	ABL32745	Human immune syste
C 865	14	10.9	7990	24	ABL32158	Human immune syste	C 938	14	10.9	16137	22	AAK87691	Human immune/haema
C 866	14	10.9	8091	23	ABL28603	Drosophila melanog	C 939	14	10.9	16279	22	AA542106	Genomic sequence #
C 867	14	10.9	8133	17	AAI28648	C. difficile toxin	C 940	14	10.9	16281	22	AAK70314	Human immune/haema
C 868	14	10.9	8133	19	AAV30560	Clostridium diffic	C 941	14	10.9	16281	23	ABK42480	Genomic sequence #
C 869	14	10.9	8283	23	ABL10886	Drosophila melanog	C 942	14	10.9	16285	22	AAK70315	Human immune/haema
C 870	14	10.9	8760	22	AAI49206	Human CHOT exons 6	C 943	14	10.9	16285	22	AAK70316	Human immune/haema
C 871	14	10.9	8920	15	AAQ62924	Carbamoyl-phosphat	C 944	14	10.9	16285	23	ABK42481	Genomic sequence #
C 872	14	10.9	8943	24	ABK39967	Human chemically p	C 945	14	10.9	16285	23	ABK42482	Genomic sequence #
C 873	14	10.9	8961	24	ABU49379	Human polynucleoti	C 946	14	10.9	16918	24	ABL33616	Human immune syste
C 874	14	10.9	8961	24	ABK28427	DNA transcription	C 947	14	10.9	17144	22	AA546665	Tumour suppressor
C 875	14	10.9	9147	23	ABL09206	Drosophila melanog	C 948	14	10.9	17173	25	ABZ80234	Mouse tramdorin 1
C 876	14	10.9	9664	22	ABA16496	Human nervous syst	C 949	14	10.9	17583	22	AAK89484	Human digestive sy
C 877	14	10.9	9664	22	AA536626	Human cardiovascular	C 950	14	10.9	18817	24	ABL70161	Chemically treated
C 878	14	10.9	9690	24	ABK52224	Human cholinergic	C 951	14	10.9	18817	24	ABL34494	Human metastasis a
C 879	14	10.9	9936	19	AAV34683	Arabidopsis thalia	C 952	14	10.9	18819	22	AAH98473	P falciparum EST-d
C 880	14	10.9	10020	24	ABL34292	Human immune syste	C 953	14	10.9	19040	22	ABA16589	Human nervous syst
C 881	14	10.9	10050	23	ABL11011	Drosophila melanog	C 954	14	10.9	19209	22	AAK70154	Human immune/haema
C 882	14	10.9	10147	23	ABL08860	Drosophila melanog	C 955	14	10.9	19209	22	AAK86103	Human immune/haema
C 883	14	10.9	10437	23	AA568713	DNA encoding novel	C 956	14	10.9	19648	23	ABL19676	Drosophila melanog
C 884	14	10.9	10983	22	ABA18199	Human nervous syst	C 957	14	10.9	19814	24	ABL70610	Chemically treated

958 14 10.9 20066 24 ABL52278 Human chemokine (C  
 959 14 10.9 21936 22 ABA15865 Human nervous syst  
 960 14 10.9 21936 22 AAL06119 Human reproductive  
 C 961 14 10.9 21936 22 AAL06758 Human reproductive  
 C 962 14 10.9 21936 22 AAL62664 Human breast or ov  
 963 14 10.9 21936 23 ABL98684 Human testicular a  
 964 14 10.9 22008 22 ABA15839 Human nervous syst  
 C 965 14 10.9 22008 22 AAS36610 Human cardiovascular  
 C 966 14 10.9 22008 25 ABZ73783 Secreted protein g  
 C 967 14 10.9 22008 25 ABT16888 Human secreted pro  
 C 968 14 10.9 22008 25 ABZ67377 Human secreted pro  
 969 14 10.9 22157 24 ABA01448 Streptococcus the  
 970 14 10.9 24413 25 ABZ74636 Secreted protein g  
 971 14 10.9 24413 25 ABT17016 Human secreted pro  
 972 14 10.9 24413 25 ABZ68156 Human secreted pro  
 973 14 10.9 24533 22 AAS27689 DNA encoding novel  
 974 14 10.9 24533 25 ABZ74637 Secreted protein g  
 975 14 10.9 24533 25 ABT17017 Human secreted pro  
 976 14 10.9 24533 25 ABZ68157 Human secreted pro  
 C 977 14 10.9 25701 22 AAL07078 Human reproductive  
 C 978 14 10.9 25758 22 AAL07077 Human reproductive  
 C 979 14 10.9 25971 22 AAK86336 Human immune/haema  
 C 980 14 10.9 25975 22 AAL04955 Human reproductive  
 C 981 14 10.9 25975 23 ABZ97849 Human testicular a  
 982 14 10.9 26767 23 ABL08654 Drosophila melanog  
 983 14 10.9 26960 23 ABL08712 Drosophila melanog  
 984 14 10.9 28588 22 AAK89418 Human digestive sy  
 985 14 10.9 28588 22 AAS31923 Human liver associ  
 986 14 10.9 28588 24 AEN90278 Human liver antige  
 C 987 14 10.9 28772 24 AAK83555 Human cDNA differe  
 988 14 10.9 28874 21 ABA81505 N. meningitidis pa  
 C 989 14 10.9 29555 23 ABL18446 Drosophila melanog  
 990 14 10.9 30549 24 AEN90859 Staphylococcus epi  
 991 14 10.9 31352 24 AAL39687 Genomic DNA encodi  
 C 992 14 10.9 31804 22 AAK90698 Human digestive sy  
 993 14 10.9 32174 22 ABA15665 Human nervous syst  
 C 994 14 10.9 32174 22 ABA19477 Human nervous syst  
 C 995 14 10.9 32174 22 ABA20359 Human nervous syst  
 996 14 10.9 32174 22 ABA21505 Human nervous syst  
 997 14 10.9 32174 22 AAL36280 Human musculoskele  
 C 998 14 10.9 32174 22 AAS32655 Human genomic DNA  
 C 999 14 10.9 32174 22 AAS34394 Human DNA for a no  
 1000 14 10.9 32174 22 AAS30638 DNA encoding novel

## ALIGNMENTS

RESULT 1  
 ID AAA71434 standard; DNA; 128 BP.  
 AC AAA71434;  
 XX  
 XX  
 DT 01-DEC-2000 (first entry)  
 XX  
 DE Human megisin promoter fragment DNA.  
 XX  
 KW Promoter; megisin; human; protein isolation; screening. ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200043528-A1.  
 XX  
 XX  
 PD 27-JUL-2000.  
 XX  
 XX  
 PF 25-JAN-2000; 2000WO-JP00350.  
 XX  
 XX  
 PR 25-JAN-1999; 99JP-0015667.  
 XX  
 XX  
 PA (KURO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.  
 XX  
 XX

PI Miyata T;  
 XX  
 DR WPI; 2000-543257/49.  
 XX  
 PT DNA for promoter region of megisin useful for screening proteins  
 XX  
 PS Claim 1; Page 32; 45pp; Japanese.  
 CC This invention describes a novel DNA sequence (I) representing a promoter  
 CC region having part or all of a specific base sequence. The invention also  
 CC describes (1) a vector containing (I); (2) a cell transformed by the  
 CC above vector; and (3) protein produced using (I). (I) is useful for  
 CC screening and isolating proteins (especially transcription factors). This  
 CC sequence represents the human megisin promoter which is described in the  
 CC method of the invention.  
 XX  
 SQ Sequence 128 BP; 44 A; 29 C; 24 G; 29 T; 2 other;  
 Query Match 100.0%; Score 128; DB 21; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-60;  
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AATGAACATACATACACCAACCTTAGTCAGATACACTTTGAAACCTGTTCAAAACCTA 60  
 Db 1 AATGAACATACATACACCAACCTTAGTCAGATACACTTTGAAACCTGTTCAAAACCTA 60  
 Qy 61 AATGCTTATAGARRCTTGAGAGACAGAGCTGCTCTGAGTCATAGGAGGAGCCATCCCA 120  
 Db 61 AATGCTTATAGARRCTTGAGAGACAGAGCTGCTCTGAGTCATAGGAGGAGCCATCCCA 120  
 Qy 121 GAAGCCAG 128  
 Db 121 GAAGCCAG 128  
 RESULT 2  
 ID AAA71435 standard; DNA; 1431 BP.  
 XX  
 AC AAA71435;  
 XX  
 DT 01-DEC-2000 (first entry)  
 XX  
 DE Human megisin promoter fragment DNA.  
 XX  
 KW Promoter; megisin; human; protein isolation; screening. ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200043528-A1.  
 XX  
 XX  
 PD 27-JUL-2000.  
 XX  
 XX  
 PF 25-JAN-2000; 2000WO-JP00350.  
 XX  
 XX  
 PR 25-JAN-1999; 99JP-0015667.  
 XX  
 XX  
 PA (KURO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.  
 XX  
 XX  
 PI Miyata T;  
 XX  
 DR WPI; 2000-543257/49.  
 XX  
 PT DNA for promoter region of megisin useful for screening proteins  
 XX  
 PS Disclosure; Fig 2; 45pp; Japanese.  
 CC This invention describes a novel DNA sequence (I) representing a promoter  
 CC region having part or all of a specific base sequence. The invention also  
 CC describes (1) a vector containing (I); (2) a cell transformed by the  
 CC above vector; and (3) protein produced using (I). (I) is useful for  
 CC screening and isolating proteins (especially transcription factors). This

CC sequence represents a fragment of the human megalin promoter which is  
CC described in the method of the invention.

XX	SQ	Sequence 1431 BP; 466 A; 266 C; 303 G; 394 T; 2 other;
	Query Match	100.0%; Score 128; DB 21; Length 1431;
	Best Local Similarity	100.0%; Pred. No. 5.4e-60;
	Matches 128; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	AATGAACCTCATAAACAACCACCTTAGTCAGATACTACTTTGAAACCTGGTTCAAAACCTA 60       
Db	1304	AATGAACCTCATAAACAACCACCTTAGTCAGATACTACTTTGAAACCTGGTTCAAAACCTA 1363       
QY	61	AATGCCTTAATGAARCTTGAGAGACAGTGTCTGTCTCAGCTCATAGGAGGCCATCCCA 120       
Db	1364	AATGCCTTAATGAARRCTTGAGAGACAGTGTCTGTCTCAGCTCATAGGAGGCCATCCCA 1423       
QY	121	GAAGCCAG 128       
Db	1424	GAAGCCAG 1431       

RESULT 3	
AAAT1448	
ID	AAA171448 standard; DNA; 30 BP.
XX	
AC	AAA171448;
XX	
DT	01-DEC-2000 (first entry)
XX	
DE	Human mesgin promoter PCR primer SEQ ID NO: 15.
XX	
KW	Promoter; mesgin; human; protein isolation; screening. PCR primer; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200043528-A1.
XX	
PD	27-JUL-2000.
XX	
FF	25-JAN-2000; 2000WO-JP00350.
XX	
PR	25-JAN-1999; 99JP-0015667.
XX	

XX	(KURO//) KUROKAWA K.
PA	(MIYA//) MIYATA T.
XX	
XX	
PI	Miyata T;
XX	
DR	WPI; 2000-543257/49.
XX	
XX	DNA for promoter region of meglin useful for screening proteins -
PT	
PT	Example 5; Page 40; 45pp; Japanese.
PS	
XX	
CC	This invention describes a novel DNA sequence (I) representing a
CC	promoter region having part or all of a specific base sequence. The
CC	invention also describes (1) a vector containing (I); (2) a cell
CC	transformed by the above vector; and (3) protein produced using (I). (I)
CC	is useful for screening and isolating proteins (especially transcription
CC	factors): AAA71434-A71469 represent PCR primers used in the method
CC	described in the invention.

XX	Sequence	30 BP;	13 A;	8 C;	3 G;	6 T;	0 other;
QY	Query Match	23.4%;	Score	30;	DB	21;	Length 30;
	Best Local Similarity	100.0%;	Pred. No.	1.6e-06;			
	Matches	30;	Conservative	0;	Mismatches	0;	Indels 0;
	Gaps	0;					
QY	1	AATGA	ACTACAT	CAACAA	CAACCTT	TAGTCAG	30
Db	1	AATGA	ACTACAT	CAACAA	CAACCTT	TAGTCAG	30

RESULT 4  
 AAA71449  
 ID AAA71449 standard; DNA; 30 BP.  
 XX  
 AC AAA71449;  
 XX  
 DT 01-DEC-2000 (first entry)  
 XX  
 DE Human megasin promoter PCR primer SEQ ID NO: 16.  
 XX  
 KW Promoter; megasin; human; protein isolation; screening. PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200043528-A1.  
 XX  
 DT 27-JUL-2000.  
 XX  
 PP 25-JAN-2000; 2000WO-JP00350.  
 XX  
 PR 25-JAN-1999; 99JP-0015667.  
 XX  
 PA (KURO/) KUROKAWA K.  
 PA (MIYA/) MIYATA T.  
 XX  
 PI Miyata T;  
 XX  
 WI 2000-543257/49.  
 DR  
 XX  
 PT DNA for promoter region of megasin useful for screening proteins -  
 XX  
 PS Example 5; Page 40; 45pp; Japanese.  
 XX  
 CC This invention describes a novel DNA sequence (I) representing a  
 CC promoter region having part or all of a specific base sequence. The  
 CC invention also describes (1) a vector containing (I); (2) a cell  
 CC transformed by the above vector; and (3) protein produced using (I). (I)  
 CC is useful for screening and isolating proteins (especially transcription  
 CC factors). AAA71434-A71469 represent PCR primers used in the method  
 CC described in the invention.

[illegible]

RESULT 5  
AAA71450  
ID AAA71450 standard; DNA; 30 BP.  
XX  
XX AC AAA71450;  
XX  
XX DT 01-DEC-2000 (first entry)  
XX  
XX DE Human meg3in promoter PCR primer SEQ ID NO: 17.  
XX  
XX KW Promoter; meg3in; human; protein isolation; screening. PCR primer; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200043528-A1.  
XX  
XX PD 27-JUL-2000.  
XX  
XX PF 25-JAN-2000; 2000WO-JP00350.  
XX

```
PR 25-JAN-1999; 99JP-0015667.
XX (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX Miyata T;
XX WPI; 2000-543257/49.
XX DNA for promoter region of megin useful for screening proteins -
XX Example 5; Page 41; 45pp; Japanese.
XX This invention describes a novel DNA sequence (I) representing a
CC promoter region having part or all of a specific base sequence. The
CC invention also describes (1) a vector containing (I); (2) a cell
CC transformed by the above vector; and (3) protein produced using (I). (I)
CC is useful for screening and isolating proteins (especially transcription
CC factors). AAA71434-A71469 represent PCR primers used in the method
CC described in the invention.
XX
XX Sequence 30 BP; 11 A; 7 C; 3 G; 9 T; 0 other;
SQ Query Match 23.4%; Score 30; DB 21; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 32 TACTACTTTGAAACCTGGTTCAAACTAA 61
Db 1 TACTACTTTGAAACCTGGTTCAAACTAA 30
RESULT 6
AAA71453
ID AAA71453 standard; DNA; 25 BP.
AC AAA71453;
XX
XX 01-DEC-2000 (first entry)
XX Human megin promoter PCR primer SEQ ID NO: 20.
XX Promoter; megin; human; protein isolation; screening. PCR primer; ss.
XX Homo sapiens.
XX WO200043528-A1.
XX 27-JUL-2000.
XX
XX (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX Miyata T;
XX WPI; 2000-543257/49.
XX DNA for promoter region of megin useful for screening proteins -
XX Example 4; Page 42; 45pp; Japanese.
XX This invention describes a novel DNA sequence (I) representing a
CC promoter region having part or all of a specific base sequence. The
CC invention also describes (1) a vector containing (I); (2) a cell
CC transformed by the above vector; and (3) protein produced using (I). (I)
CC is useful for screening and isolating proteins (especially transcription
CC factors). AAA71434-A71469 represent PCR primers used in the method
CC described in the invention.
XX
```

```
SQ Sequence 25 BP; 8 A; 7 C; 2 G; 8 T; 0 other;
Query Match 19.5%; Score 25; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 16 AACACCTTACTGACATACTACTTT 40
Db 1 AACACCTTACTGACATACTACTTT 25
RESULT 7
AAA71452
ID AAA71452 standard; DNA; 23 BP.
XX
XX AAA71452;
XX
XX 01-DEC-2000 (first entry)
XX Human megin promoter PCR primer SEQ ID NO: 19.
XX Promoter; megin; human; protein isolation; screening. PCR primer; ss.
XX Homo sapiens.
XX WO200043528-A1.
XX 27-JUL-2000.
XX
XX 25-JAN-2000; 2000WO-JP00350.
XX
XX 25-JAN-1999; 99JP-0015667.
XX (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX Miyata T;
XX WPI; 2000-543257/49.
XX DNA for promoter region of megin useful for screening proteins -
XX Example 4; Page 42; 45pp; Japanese.
XX This invention describes a novel DNA sequence (I) representing a
CC promoter region having part or all of a specific base sequence. The
CC invention also describes (1) a vector containing (I); (2) a cell
CC transformed by the above vector; and (3) protein produced using (I). (I)
CC is useful for screening and isolating proteins (especially transcription
CC factors). AAA71434-A71469 represent PCR primers used in the method
CC described in the invention.
XX
XX Sequence 23 BP; 11 A; 7 C; 2 G; 3 T; 0 other;
Query Match 17.2%; Score 22; DB 21; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 AATGAACATACATACACACC 22
Db 2 AATGAACATACATACACACC 23
RESULT 8
ABL34015/c
ID ABL34015 standard; DNA; 6182 BP.
XX
XX ABL34015;
XX
XX 26-MAR-2002 (first entry)
XX Human immune system associated gene SEQ ID NO: 1988.
XX
```



CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms tumour.

SQ Sequence 482 BP; 180 A; 79 C; 93 G; 123 T; 7 other;

Query Match 14.8%; Score 19; DB 24; Length 482;

Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 AAAACCTAAAGCTTATAA 71

Db 41 AAAACCTAAAGCTTATAA 59

RESULT 10

AAC39949/c

ID AAC39949 standard; DNA; 1523 BP.

XX AC AAC39949;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26493.

XX KW Hybridisation assay; genetic mapping; gene expression control;

XX KW protein identification; signal transduction pathway;

XX KW metabolic pathway; promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136382.

PR 01-JUN-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 05-AUG-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 23-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142154.

PR 05-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 14.8%; Score 19; DB 21; Length 1523;  
 Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 15 CAACACCTTAGTCAGATA 33  
 Db 646 CAACACCTTAGTCAGATA 628  
 RESULT 11  
 AAZ69316  
 ID AAZ69316 standard; DNA; 47 BP.  
 XX  
 AC AAZ69316;  
 XX  
 DT 10-SEP-2001 (first entry)  
 XX  
 DE Human map-related biallelic marker SEQ ID NO:3672.  
 XX  
 KW Human genome; biallelic marker; high density disequilibrium map;  
 KW genomic map; haplotype; phenotype; polymorphic base; genotyping;  
 KW haplotyping; hybridisation; identification; characterisation;  
 KW diagnosis; single nucleotide polymorphism; SNP; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT variation replace(24,C)  
 FT /\*tag= a  
 FT /standard\_name= "single nucleotide polymorphism"  
 XX  
 EN WO9954500-A2.  
 XX  
 PD 28-OCT-1999.  
 XX  
 PF 21-APR-1999; 99WO-IB00822.  
 XX  
 PR 21-APR-1998; 98US-0082614.  
 PR 23-NOV-1998; 98US-0109732.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Cohen D, Blumenfeld M, Chumakov I;  
 XX  
 DR WPI; 2000-013267/01.  
 XX  
 PT Novel biallelic markers used to construct a high density disequilibrium  
 map of the human genome  
 XX  
 PS Claim 3; Page 1014; 2745pp; English.  
 XX  
 CC AAZ65654 to AAZ69578 represent human biallelic markers from the present  
 invention, which contain a polymorphic base at position 24 of their  
 nucleotide sequences. AAZ69579 to AAZ77440 represent amplification  
 primers for the biallelic markers. The biallelic markers of the  
 mapping of the human genome, and in complex association studies and  
 haplotyping studies which are useful in determining the genetic basis  
 for disease states. Compositions and methods of the invention can also  
 be useful for the identification of the targets for the development of  
 pharmaceutical agents and diagnostic methods, as well as the  
 characterisation of the differential efficacious responses to and side  
 effects from pharmaceutical agents acting on a disease as well as other  
 treatment.  
 CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297  
 and 3367, are not actually given a sequence in the Sequence Listing  
 from the present invention.  
 CC  
 XX Sequence 47 BP; 15 A; 14 C; 11 G; 7 T; 0 other;  
 Query Match 14.1%; Score 18; DB 21; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 104 ATAGGGAAGCCATCCAC 121



Db 23 ATAGGAGCCATCCAG 40  
|||||

RESULT 12  
AAS83451/C  
ID AAS83451 standard; cDNA; 1227 BP.

XX AC AAS83451;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #19255.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Dmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG19264.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity -

XX PS Claim 1; SEQ ID No 19255; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and  
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX CC and gene mapping, and in recombinant production of (II). The  
XX CC polynucleotides are also used in diagnostics as expressed sequence tags  
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques  
XX CC to restore normal activity of (II) or to treat disease states involving  
XX CC (II). (II) is useful for generating antibodies against it, detecting or  
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as  
XX CC a food supplement. (II) and its binding partners are useful in medical  
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating  
XX CC disorders involving aberrant protein expression or biological activity.  
XX CC The polypeptide and polynucleotide sequences have applications in  
XX CC diagnostics, forensics, gene mapping, identification of mutations  
XX CC responsible for genetic disorders or other traits to assess biodiversity  
XX CC and to produce other types of data and products dependent on DNA and  
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human  
XX CC diagnostic coding sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 1227 BP; 314 A; 290 C; 326 G; 295 T; 2 other;

Query Match 14.1%; Score 18; DB 23; Length 1227;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 107 GGGAGCCATCCAGAG 124  
|||||

Db 545 GGGAGCCATCCAGAG 528  
|||||

RESULT 13

ABT11173

ID ABT11173 standard; DNA; 168174 BP.

XX AC ABT11173;

XX DT 05-DEC-2002 (first entry)

XX DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 63.

XX KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;  
XX KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;  
XX KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;  
XX KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;  
XX KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;  
XX KW ds.

XX OS Homo sapiens.

XX FN WO200262825-A2.

XX PD 15-AUG-2002.

XX PF 07-FEB-2002; 2002WO-US03546.

XX PR 08-FEB-2001; 2001US-267515P.

XX PR 21-AUG-2001; 2001US-314248P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Barnes G, Meyer J;

XX DR WPI; 2002-627522/67.

XX PT New isolated nucleic acid molecule with an allelic variant of a  
XX PT polymorphic region of an 5-LO gene, useful for diagnosing and/or  
XX PT prognosticating disorders associated with an aberrant inflammatory  
XX PT response such as asthma -

XX PS Disclosure; Fig 4; 290pp; English.

XX CC The invention relates to an isolated human nucleic acid molecule  
XX CC comprising an allelic variant of a polymorphic region of a 5-lipoxygenase  
XX CC (5-LO) gene, where the allelic variant comprises one or more nucleotide  
XX CC selected from any of 3, 20 or 21 base pair sequences, given in the  
XX CC specification, or their complement. The compositions and methods of the  
XX CC present invention are useful for diagnosing and/or prognosticating disorders  
XX CC associated with an aberrant inflammatory response such as asthma,  
XX CC bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,  
XX CC rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic  
XX CC rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory  
XX CC disease, atopic and contact dermatitis. The nucleic acid molecules can  
XX CC also be useful for identifying an individual amongst other individuals  
XX CC from the same species for use in forensic medicine and paternity testing.  
XX CC This polynucleotide sequence represents DNA relating to the human 5-  
XX CC lipoxygenase (5-LO) gene of the invention.

SQ Sequence 168174 BP; 46808 A; 36942 C; 36942 G; 46474 T; 1508 other;  
Query Match 14.1%; Score 18; DB 24; Length 168174;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 TGCTCTGAGTCATAGGA 110  
|||||

Db 49803 TGCTCTGAGTCATAGGA 49820  
|||||

RESULT 14

ABT11114

ID ABT11114 standard; DNA; 168273 BP.

XX AC AET11114;  
 XX DT 05-DEC-2002 (first entry)  
 XX DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 2.  
 XX KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;  
 KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;  
 KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;  
 KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;  
 KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;  
 KW ds.  
 XX OS Homo sapiens.  
 XX PN WO200262825-A2.  
 XX PD 15-AUG-2002.  
 XX PF 07-FEB-2002; 2002WO-US03546.  
 XX PR 08-FEB-2001; 2001US-267515P.  
 XX PR 21-AUG-2001; 2001US-314248P.  
 XX PA (MILL-) MILLENNIUM PHARM INC.  
 XX PI Barnes G, Meyer J;  
 XX PI WPI; 2002-627522/67.  
 XX PT New isolated nucleic acid molecule with an allelic variant of a  
 PT polymorphic region of an 5-LO gene, useful for diagnosing and/or  
 PT prognosticating disorders associated with an aberrant inflammatory  
 PT response such as asthma  
 XX PS Disclosure; Fig 2; 290pp; English.  
 XX CC The invention relates to an isolated human nucleic acid molecule  
 CC comprising an allelic variant of a polymorphic region of a 5-lipoxygenase  
 CC (5-LO) gene, where the allelic variant comprises one or more nucleotide  
 CC selected from any of 3, 20 or 21 base pair sequences, given in the  
 CC specification, or their complement. The compositions and methods of the  
 CC present invention are useful for diagnosing and/or prognosing disorders  
 CC associated with an aberrant inflammatory response such as asthma,  
 CC bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,  
 CC rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic  
 CC rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory  
 CC disease, atopic and contact dermatitis. The nucleic acid molecules can  
 CC also be useful for identifying an individual amongst other individuals  
 CC from the same species for use in forensic medicine and paternity testing.  
 CC This polynucleotide sequence represents DNA relating to the human 5-  
 CC lipoxygenase (5-LO) gene of the invention.  
 XX SQ Sequence 168273 BP; 46934 A; 36467 C; 36966 G; 46498 T; 1508 other;  
 Query Match 14.1%; Score 18; DB 24; Length 168273;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 93 TGCTCTGAGTCATAGGGA 110  
 |||||  
 Db 49853 TGCTCTGAGTCATAGGGA 49870  
 RESULT 15  
 AAD36511/c  
 ID AAD36511 standard; DNA; 169998 BP.  
 XX AC AAD36511;  
 XX DT 09-AUG-2002 (first entry)  
 XX

DE Human Her-1 gene.  
 XX KW Human; epidermal growth factor receptor; hyperproliferative disease;  
 KW Her1; prophylaxis; psoriasis; tumour; cancer; gene; ds.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT exon 1208..1472  
 FT /tag= a  
 FT intron 1473..124390  
 FT /tag= b  
 FT exon 124391..124544  
 FT /tag= c  
 FT intron 124545..125409  
 FT /tag= d  
 FT exon 125410..125595  
 FT /tag= e  
 FT intron 125596..128711  
 FT /tag= f  
 FT exon 128712..128848  
 FT /tag= g  
 FT intron 128849..133400  
 FT /tag= h  
 FT exon 133401..133469  
 FT /tag= i  
 FT intron 133470..134652  
 FT /tag= j  
 FT exon 134653..134773  
 FT /tag= k  
 FT intron 134774..136116  
 FT /tag= l  
 FT exon 136117..136261  
 FT /tag= m  
 FT intron 136262..137936  
 FT /tag= n  
 FT exon 137937..138053  
 FT /tag= o  
 FT intron 138054..138637  
 FT /tag= p  
 FT exon 138638..138766  
 FT /tag= q  
 FT intron 138767..138864  
 FT /tag= r  
 FT exon 138865..138940  
 FT /tag= s  
 FT intron 138941..139765  
 FT /tag= t  
 FT exon 139766..139860  
 FT /tag= u  
 FT intron 139861..142245  
 FT /tag= v  
 FT exon 142246..142445  
 FT /tag= w  
 FT intron 142446..143605  
 FT /tag= x  
 FT exon 143606..143738  
 FT /tag= y  
 FT intron 143739..145838  
 FT /tag= z  
 FT exon 145839..145931  
 FT /tag= aa  
 FT intron 145932..147385  
 FT /tag= ab  
 FT exon 147386..147544  
 FT /tag= ac  
 FT intron 147545..153274  
 FT /tag= ad  
 FT exon 153275..153321  
 FT /tag= ae  
 FT intron 153322..155088  
 FT /tag= af  
 FT exon 155089..155231

```

PT intron /tag= ag
PT 155232..156025
PT /tag= ah
FT exon 156026..156151
FT /tag= ai
FT intron 156152..156826
FT /tag= aj
FT exon 156827..156928
FT /tag= ak
FT intron 156929..163399
FT /tag= al
FT exon 163400..163586
FT /tag= am
XX
XX WO200226758-A1.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US30551.
XX
XX 29-SEP-2000; 2000US-0676610.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Bennett CF, Wyatt JR, Freier SM;
XX
XX WPI; 2002-394234/42.
XX
XX Novel antisense oligonucleotide that specifically hybridizes with and
XX inhibits nucleic acid encoding epidermal growth factor receptor, useful
XX for treating hyperproliferative disease such as cancer or psoriasis -
XX
XX Example 19; Page 67-121; 169pp; English.
XX
XX The invention relates to an antisense oligonucleotide targetted to a
XX nucleic acid molecule encoding human epidermal growth factor receptor
XX (Her1) to inhibit its expression. The antisense compounds are useful
XX for treating diseases or conditions associated with Her-1 such as
XX hyperproliferative diseases especially cancer (lung, ovarian, colon
XX or prostate cancer) and psoriasis. They are also useful as research
XX reagents, diagnostics, therapeutics, kits and prophylactically e.g.
XX to prevent or delay tumour formation. The present sequence is
XX human Her-1 gene.
XX
XX Sequence 169998 BP; 46143 A; 38164 C; 37751 G; 47940 T; 0 other;
XX
XX Query Match 14.1%; Score 18; DB 24; Length 169998;
XX Best Local Similarity 100.0%; Pred. No. 7.4;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 77 TTGAGAGACAGTGTGTG 94
XX |||||
XX Db 12273 TTGAGAGACAGTGTGTG 12256
XX
XX Search completed: August 14, 2003, 18:44:04
XX Job time : 227 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2003, 18:36:54 ; Search time 1668 Seconds  
(without alignments)

1865.092 Million cell updates/sec

Title: US-09-889-611-1

Perfect score: 128

Sequence: 1 aatgaactacatacaacca.....gaagcattccagagcg 128

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estnu.\*

5: em\_estov.\*

6: em\_estpi.\*

7: em\_estro.\*

8: em\_hic.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hic.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_nam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	23.4	300	9	AU100197
2	19	14.8	142	9	AA975162
3	19	14.8	294	10	BF365378
4	19	14.8	315	14	CA412191

78	17	13.3	359	28	BZ167901	BZ167901 CH230-514	c 151	17	13.3	957	13	EX349544	BX349544
79	17	13.3	360	14	C48938	C48938 C48938 Yui1	c 152	17	13.3	1005	10	EC398359	EC398359 602439826
80	17	13.3	360	14	C70159	C70159 C70159 Yui1	c 153	17	13.3	1012	13	EX446254	EX446254 EX446254
81	17	13.3	369	14	C3595434	C3595434 C864009.Y	c 154	17	13.3	1015	14	CD517260	AGENCOURT
82	17	13.3	408	28	AO983222	AO983222 RPI-23-3	c 155	17	13.3	1036	10	BE393345	BE393345 601311011
83	17	13.3	423	13	BY275685	BY275685 BY275685	c 156	17	13.3	1037	13	BQ481854	BQ481854 PV G8a001
84	17	13.3	435	9	AI184708	AI184708 qd86d11.x	c 157	17	13.3	1044	13	BQ072923	AGENCOURT
85	17	13.3	484	28	AO471300	AO471300 CITE1-E1	c 158	17	13.3	1065	13	BQ073024	AGENCOURT
86	17	13.3	504	13	BW198546	BW198546 BW198546	c 159	17	13.3	1082	12	EN913215	AGENCOURT
87	17	13.3	513	14	CB387336	CB387336 OSTF078G1	c 160	17	13.3	1082	13	BQ063566	AGENCOURT
88	17	13.3	531	28	AQ080964	AQ080964 HS 3155.A	c 161	17	13.3	1086	14	CD242377	AGENCOURT
89	17	13.3	532	14	CB406497	CB406497 OSTR073F1	c 162	17	13.3	1096	12	EN456786	AGENCOURT
90	17	13.3	533	28	AZ484211	AZ484211 IM0310104	c 163	17	13.3	1101	29	CNS0510V	Tetradon
91	17	13.3	537	28	B76574	B76574 RPI11-15M2	c 164	17	13.3	1116	29	CC220756	CH261-83M
92	17	13.3	550	14	CB387018	CB387018 OSTF073F1	c 165	17	13.3	1154	12	EN913780	AGENCOURT
93	17	13.3	556	28	AZ652048	AZ652048 IM0525B07	c 166	17	13.3	1156	29	CC203826	CH261-31M
94	17	13.3	562	10	BZ213741	BZ213741 BZ213741	c 167	17	13.3	1171	12	EN906884	AGENCOURT
95	17	13.3	567	28	AZ369484	AZ369484 IM0120P05	c 168	17	13.3	1186	12	EN466249	AGENCOURT
96	17	13.3	569	28	CC162058	CC162058 i181504.b	c 169	17	13.3	1201	13	BX355636	BX355636
97	17	13.3	608	28	B83564	B83564 RPI11-15M2	c 170	17	13.3	1201	13	BX396110	BX396110
98	17	13.3	623	13	BU077232	BU077232 i18G05.Y	c 171	17	13.3	1201	13	EX462140	EX462140
99	17	13.3	627	12	B1771497	B1771497 603095Y4	c 172	17	13.3	1201	13	EX462141	EX462141
100	17	13.3	632	13	BU490464	BU490464 604129225	c 173	17	13.3	1210	13	BQ944174	AGENCOURT
101	17	13.3	633	14	CAY30610	CAY30610 wp1C.pk0	c 174	17	13.3	1788	11	AK040675	Mus muscu
102	17	13.3	646	29	AG133432	AG133432 Pan trogl	c 175	17	13.3	2112	11	AK086117	Mus muscu
103	17	13.3	646	29	AG133432	AG133432 Pan trogl	c 176	17	13.3	3336	11	AK048821	Mus muscu
104	17	13.3	652	14	CB103358	CB103358 ADP SQ006	c 177	17	13.3	3764	11	AK087514	Mus muscu
105	17	13.3	666	14	B755815	B755815 B755815	c 178	17	13.3	3994	11	AK082641	Mus muscu
106	17	13.3	666	29	AG035135	AG035135 Pan trogl	c 179	16	12.5	191	28	AZ270256	Mus muscu
107	17	13.3	667	13	BQ997689	BQ997689 QG617C07.	c 180	16	12.5	199	28	AQ389352	RPCI11-14
108	17	13.3	675	10	BG503318	BG503318 602550729	c 181	16	12.5	205	9	AL706056	DKF2p686L
109	17	13.3	675	10	BE513780	BE513780 601315603	c 182	16	12.5	205	9	BX331806	BX331806
110	17	13.3	677	28	AZ729438	AZ729438 RPI-24-1	c 183	16	12.5	221	9	AA635816	nr38a07.s
111	17	13.3	691	12	B1911265	B1911265 603065816	c 184	16	12.5	222	9	AA570634	AA570634 nj08C04.s
112	17	13.3	695	13	BQ874693	BQ874693 QG16A13.Y	c 185	16	12.5	222	14	F04100	F04100 HSC2JF042.n
113	17	13.3	695	28	BZ024104	BZ024104 oed1A12.	c 186	16	12.5	222	9	AI625102	ts47a01.x
114	17	13.3	698	12	B1835122	B1835122 603087787	c 187	16	12.5	228	9	AV319739	AV319739
115	17	13.3	709	14	CA503672	CA503672 UI-R-FS1-	c 188	16	12.5	228	28	BH818691	BH818691 BACP11-F
116	17	13.3	713	13	BX371615	BX371615 BX371615	c 189	16	12.5	239	9	AA778804	AA778804 zj45h06.s
117	17	13.3	729	28	AZ311916	AZ311916 IM0027J03	c 190	16	12.5	241	12	B1074433	B1074433 B1074433
118	17	13.3	746	29	BZ633017	BZ633017 PUAB31TB	c 191	16	12.5	251	14	N74310	N74310 za78c01.s1
119	17	13.3	748	14	CA504060	CA504060 UI-R-FJ0-	c 192	16	12.5	252	14	H92750	H92750 yt92e01.s1
120	17	13.3	751	12	B1837822	B1837822 603083527	c 193	16	12.5	254	14	Z40943	Z40943 HSC3JG042.n
121	17	13.3	754	28	AO856227	AO856227 nbe00001H	c 194	16	12.5	259	10	BF067189	BF067189 st34a12.y
122	17	13.3	768	29	BX136496	BX136496 Dario rex	c 195	16	12.5	262	9	AA598989	AA598989 ae34e09.s
123	17	13.3	777	12	B1521086	B1521086 603081679	c 196	16	12.5	262	10	B791485	B791485 B791485
124	17	13.3	777	12	B1835945	B1835945 603085690	c 197	16	12.5	263	9	AW481932	AW481932 39398.MAR
125	17	13.3	795	13	BU929769	BU929769 AGENCOURT	c 198	16	12.5	266	10	BF329550	BF329550 CM0-BN026
126	17	13.3	799	13	BQ433411	BQ433411 AGENCOURT	c 199	16	12.5	266	10	B844446	B844446 B844446
127	17	13.3	807	12	B1837038	B1837038 603084587	c 200	16	12.5	269	28	B31346	B31346 HS-1008-A2-
128	17	13.3	818	29	BZ633020	BZ633020 PUAB31TD	c 201	16	12.5	271	10	BG007056	BG007056 RCI-GN023
129	17	13.3	821	10	BG684951	BG684951 602636690	c 202	16	12.5	275	10	BG165126	BG165126 602343950
130	17	13.3	831	29	CC116458	CC116458 NDL.53J06.	c 203	16	12.5	279	9	AV224778	AV224778 AV224778
131	17	13.3	846	14	CA767248	CA767248 AF53-Rpf	c 204	16	12.5	289	10	B8030448	B8030448 B8030448
132	17	13.3	849	12	B1755764	B1755764 603025086	c 205	16	12.5	290	9	AV057682	AV057682 AV057682
133	17	13.3	855	12	B1819859	B1819859 603041392	c 206	16	12.5	299	9	AV114601	AV114601 AV114601
134	17	13.3	856	13	BU447330	BU447330 603766378	c 207	16	12.5	299	14	T24084	seqQ2272.b4H
135	17	13.3	858	12	B1909237	B1909237 603070278	c 208	16	12.5	300	9	AV169199	AV169199 AV169199
136	17	13.3	860	28	BZ174322	BZ174322 CH230-442	c 209	16	12.5	302	9	AA991980	AA991980 CT01C03.s
137	17	13.3	876	28	BZ152512	BZ152512 CH230-346	c 210	16	12.5	316	14	T33041	T33041 EST56443.Hu
138	17	13.3	878	12	B1906110	B1906110 603062510	c 211	16	12.5	317	10	BE275640	BE275640 BE275640
139	17	13.3	886	10	BF571053	BF571053 602075986	c 212	16	12.5	317	10	BE120288	BE120288 UI-R-CAO-
140	17	13.3	903	12	B1834149	B1834149 603085132	c 213	16	12.5	329	28	AO116404	AO116404 CIT978SK-
141	17	13.3	916	12	B1834138	B1834138 603085116	c 214	16	12.5	331	9	AN582096	AN582096 NR3-ST019
142	17	13.3	916	13	EX428635	EX428635 EX428635	c 215	16	12.5	347	10	B8667582	B8667582 B8667582
143	17	13.3	926	10	BG493505	BG493505 602542411	c 216	16	12.5	347	28	BH439214	BH439214 BOR0987TF
144	17	13.3	930	13	BQ465113	BQ465113 EX465113	c 217	16	12.5	348	9	AA875859	AA875859 OB34d01.s
145	17	13.3	932	13	BQ706513	BQ706513 AGENCOURT	c 218	16	12.5	352	10	BF336089	BF336089 OV4-CT049
146	17	13.3	938	10	BE563703	BE563703 601335323	c 219	16	12.5	352	13	B0719956	B0719956 AGENCOURT
147	17	13.3	948	29	AN505881	AN505881 Tetradon	c 220	16	12.5	357	14	H32744	H32744 yt92c01.s1
148	17	13.3	949	10	BF582098	BF582098 602099283	c 221	16	12.5	359	14	H22974	H22974 YMS5B05.s1
149	17	13.3	951	12	B1832805	B1832805 603082615	c 222	16	12.5	360	9	AV201979	AV201979 AV201979
150	17	13.3	953	13	BQ708648	BQ708648 AGENCOURT	c 223	16	12.5	360	12	BI508674	BI508674 BI508674

c 224	16	12.5	365	14	T74904	T74904 yc58g07.s1	c 297	16	12.5	508	28	BH825937
c 225	16	12.5	370	9	AW571946	AW571946 xg1ld12.x	c 298	16	12.5	511	28	AQ369308
c 226	16	12.5	372	28	BH040172	BH040172 RPCI-24-2	c 299	16	12.5	512	9	AU149852
c 227	16	12.5	376	9	AI207849	AI207849 an06a05.x	c 300	16	12.5	512	13	BQ818619
c 228	16	12.5	381	10	BF373680	BF373680 MR3-SN001	c 301	16	12.5	513	9	AL042177
c 229	16	12.5	382	9	AA356443	AA356443 ES164995	c 302	16	12.5	516	14	T78595
c 230	16	12.5	383	9	AA991938	AA991938 os99c03.s	c 303	16	12.5	517	10	BE190822
c 231	16	12.5	385	9	AA620520	AA620520 ae0oe10.s	c 304	16	12.5	517	28	BZ032288
c 232	16	12.5	385	12	BQ7049364	BQ7049364 BJ049364	c 305	16	12.5	518	12	BM269742
c 233	16	12.5	388	9	AI671652	AI671652 ty26c06.x	c 306	16	12.5	518	28	AZ705090
c 234	16	12.5	390	28	BZ180141	BZ180141 CH230-338	c 307	16	12.5	518	28	AQ414838
c 235	16	12.5	391	9	AW324243	AW324243 73998 MAR	c 308	16	12.5	518	28	AQ414838
c 236	16	12.5	393	13	BY634732	BY634732 BY634732	c 309	16	12.5	533	14	CA936872
c 237	16	12.5	397	9	AI885408	AI885408 w193d10.x	c 310	16	12.5	535	12	BM524907
c 238	16	12.5	402	9	AW811058	AW811058 AV811058	c 311	16	12.5	536	12	BQ704858
c 239	16	12.5	403	14	W99070	W99070 zh71h03.s1	c 312	16	12.5	538	28	BH677819
c 240	16	12.5	403	28	AQ755721	AQ755721 HS 3120.B	c 313	16	12.5	540	9	AA216577
c 241	16	12.5	404	28	AQ218781	AQ218781 HS-2190.A	c 314	16	12.5	540	9	AI959401
c 242	16	12.5	410	9	AA876345	AA876345 o124a10.s	c 315	16	12.5	542	10	BS681767
c 243	16	12.5	411	10	BG234342	BG234342 daa40c06.	c 316	16	12.5	542	12	BQ92761
c 244	16	12.5	415	10	BF589149	BF589149 nab27h02.	c 317	16	12.5	542	28	AZ594298
c 245	16	12.5	416	14	CA386931	CA386931 668666 NC	c 318	16	12.5	543	13	BQ860966
c 246	16	12.5	417	9	AI017652	AI017652 ou99b07.x	c 319	16	12.5	544	12	BZ257499
c 247	16	12.5	419	9	AU224092	AU224092 AU224092	c 320	16	12.5	544	28	AQ993830
c 248	16	12.5	420	9	AW130555	AW130555 x152h12.x	c 321	16	12.5	546	9	AU154569
c 249	16	12.5	421	29	BZ934172	BZ934172 CH240.89D	c 322	16	12.5	548	13	BY478416
c 250	16	12.5	422	10	BH821942	BH821942 BH821942	c 323	16	12.5	551	12	BQ7057402
c 251	16	12.5	423	10	BF926136	BF926136 CM2-NT019	c 324	16	12.5	551	12	BQ82061
c 252	16	12.5	423	28	AO409348	AO409348 HS 5066.B	c 325	16	12.5	551	13	BQ833570
c 253	16	12.5	424	14	H99159	H99159 yx19504.s1	c 326	16	12.5	552	12	BM106540
c 254	16	12.5	427	9	AA791133	AA791133 v91c109.x	c 327	16	12.5	553	9	AW813463
c 255	16	12.5	427	13	BQ860230	BQ860230 QGCLSD03.	c 328	16	12.5	555	9	AA533088
c 256	16	12.5	430	13	BY569049	BY569049 BY569049	c 329	16	12.5	555	10	BH760233
c 257	16	12.5	431	9	AW753521	AW753521 PMO-CT026	c 330	16	12.5	555	28	BH504338
c 258	16	12.5	431	28	AZ269028	AZ269028 RPCI-23-1	c 331	16	12.5	556	28	BH315569
c 259	16	12.5	436	9	AW598133	AW598133 sj41g10.y	c 332	16	12.5	557	14	CB500469
c 260	16	12.5	436	9	AW598133	AW598133 sj41g10.y	c 333	16	12.5	560	28	BI4456
c 261	16	12.5	436	13	BY501056	BY501056 BY501056	c 334	16	12.5	561	28	AZ510056
c 262	16	12.5	436	29	BZ650662	BZ650662 OGANH50TM	c 335	16	12.5	563	12	BI058230
c 263	16	12.5	438	28	AQ770136	AQ770136 HS 2002.A	c 336	16	12.5	564	10	BE347205
c 264	16	12.5	440	28	AQ044508	AQ044508 CIT-HSP-2	c 337	16	12.5	564	28	AZ24016
c 265	16	12.5	441	9	AI394515	AI394515 tf17a03.x	c 338	16	12.5	564	28	BH509179
c 266	16	12.5	445	14	T03650	T03650 IB671 infan	c 339	16	12.5	567	14	CB104942
c 267	16	12.5	445	28	AQ816996	AQ816996 HS 5450.B	c 340	16	12.5	569	9	AI494915
c 268	16	12.5	445	29	BZ475394	BZ475394 BQNLQ38TR	c 341	16	12.5	572	12	BG805120
c 269	16	12.5	446	9	AU223962	AU223962 AU223962	c 342	16	12.5	572	28	BH510845
c 270	16	12.5	450	9	AA134738	AA134738 z130f10.s	c 343	16	12.5	574	12	BI701328
c 271	16	12.5	451	9	AI024740	AI024740 ov76f09.x	c 344	16	12.5	577	9	AW703917
c 272	16	12.5	455	28	AQ584912	AQ584912 RPCI-11-4	c 345	16	12.5	577	14	CB3303
c 273	16	12.5	460	29	FR0041318	FR0041318 Fugu rubr	c 346	16	12.5	578	13	BQ630039
c 274	16	12.5	462	28	AQ789253	AQ789253 HS 3250.A	c 347	16	12.5	579	13	BW224602
c 275	16	12.5	463	9	AW169332	AW169332 xj22d06.x	c 348	16	12.5	579	14	CD344288
c 276	16	12.5	465	28	BH014374	BH014374 TQGBL02TH	c 349	16	12.5	580	10	BG303257
c 277	16	12.5	472	12	BZ222255	BZ222255 kx18f06.y	c 350	16	12.5	580	13	BQ080887
c 278	16	12.5	482	28	AZ784145	AZ784145 2M0026F11	c 351	16	12.5	582	9	AU151695
c 279	16	12.5	484	10	BQ041997	BQ041997 saa44b06.	c 352	16	12.5	583	9	AL706046
c 280	16	12.5	485	10	BF425038	BF425038 su53e05.y	c 353	16	12.5	583	13	BQ611736
c 281	16	12.5	487	10	BF336108	BF336108 QV4-CT049	c 354	16	12.5	584	13	BQ825533
c 282	16	12.5	487	28	BZ189077	BZ189077 CH230-247	c 355	16	12.5	585	10	BQ666386
c 283	16	12.5	488	29	BZ901022	BZ901022 CH240.20A	c 356	16	12.5	585	28	AZ942270
c 284	16	12.5	490	9	AI763434	AI763434 wh92d05.x	c 357	16	12.5	590	10	BE661460
c 285	16	12.5	494	9	AA782576	AA782576 ai57d05.s	c 358	16	12.5	590	10	BE661460
c 286	16	12.5	496	9	AA58761	AA58761 zk65d05.r	c 359	16	12.5	591	13	BQ741976
c 287	16	12.5	496	10	BF602221	BF602221 267400.WA	c 360	16	12.5	591	14	CA428971
c 288	16	12.5	496	28	AQ966026	AQ966026 LBRIG73TF	c 361	16	12.5	593	28	BH670906
c 289	16	12.5	496	28	BH5891	BH5891 HS-1062-BI-	c 362	16	12.5	593	29	CC059513
c 290	16	12.5	498	28	BH000728	BH000728 2M0288H13	c 363	16	12.5	599	10	AW936382
c 291	16	12.5	499	13	BY490636	BY490636 BY490636	c 364	16	12.5	600	12	BI986683
c 292	16	12.5	501	28	BH523556	BH523556 BQGYH69TR	c 365	16	12.5	602	9	AL715064
c 293	16	12.5	502	13	BQ082437	BQ082437 K-EST0109	c 366	16	12.5	603	28	BH448547
c 294	16	12.5	505	9	AI743925	AI743925 w934f04.x	c 367	16	12.5	605	12	BI434419
c 295	16	12.5	505	28	AQ389348	AQ389348 RPCI11-14	c 368	16	12.5	606	9	AL795657
c 296	16	12.5	508	28	BH716053	BH716053 BONGF11TF	c 369	16	12.5	606	12	BI050795
	16	12.5	508	28	BH716053	BH716053 BONGF11TF	c 369	16	12.5	607	12	BQ016158

BH825937	BACPP22-L
AQ369308	HS 5037.A
AU149852	AU149852
BQ818619	1030072A1
AL042177	DKFZp434F
T78595	y469D09.s1
BE190822	sn79508.y
BZ032288	oe09c04.
BM269742	sak19f08.
AZ705090	RPCI-23-2
AQ414838	RPCI-11-2
CA936872	say26a05.
BM524907	sal25c07.
BJ048458	BJ048458
BH677819	BMOMK025TR
AA216577	zq94c08.x
AI959401	AI959401
AI090697	qa71c09.x
BE681767	179542.WA
BJ092761	BJ092761
AZ594298	1M0406K07
BQ860966	QGC17C17
BZ257499	602967619
AQ993830	RPCI-23-3
AU154569	AU154569
BM106540	510331.MA
AA533088	nj19f10.s
BH760233	BH760233
BH504338	BQGYO06TR
CB500469	sealra022
BI4456	A-861H12.TP
AZ510056	1M0354118
BE347205	sp36c03.y
BI058230	PMO-GN049
AZ24016	1M0045G13
BH509179	BOHFT38TF
CB104942	K-EST0079
AI494915	sa92e11.y
BG805120	0612-66.M
BH510845	BOGL42F3F
AW703917	sk25g03.y
CB3303	C93303.D1ct
BQ630039	sap96f03.
BW224602	BW224602
CD344288	Rt-ESTef79
BG303257	f150806.x
BQ080887	saq31c07.
AU151695	AU151695
AL706046	DKFZp686K
BQ611736	sap64e11.
BI082553	saq34a08.
BQ666386	HS042D06-
AZ942270	2M0202A22
BE661460	62.Gnax3C
BQ741976	saq21c03.
CA428971	UI-H-FH1-
BH670906	BONGH56TR
CC059513	1124c12.D
AW936382	QV4-DT002
BI986683	3174-94.M
AL715064	AL715064
BH448547	BOHGO19TR
BI434419	EST537180
AL795657	AL795657
BI050795	MR4-GN019
BQ016158	UI-H-DT1-

C 370	16	12.5	608	10	BB551270	BB551270	BB551270	BB551270	443	16	12.5	699	14	CA438235	UI-H-DT1
C 371	16	12.5	608	13	BW202526	BW202526	BW202526	BW202526	444	16	12.5	703	28	BZ073296	lkf66b05
C 372	16	12.5	608	18	BH354983	CH230-55E	BH354983	CH230-55E	445	16	12.5	704	14	CA058112	ssalqta5
C 373	16	12.5	609	9	AW582072	NR3-ST019	AW582072	NR3-ST019	446	16	12.5	707	29	BX246898	Danio rer
C 374	16	12.5	609	10	BE610536	sq77f07.Y	BE610536	sq77f07.Y	C 447	16	12.5	707	28	BZ013739	oe86904.
C 375	16	12.5	609	12	BQ017198	UI-H-DT1	BQ017198	UI-H-DT1	C 448	16	12.5	708	29	AG160074	Pan trogl
C 376	16	12.5	609	13	CA776396	UI-H-FE0	CA776396	UI-H-FE0	449	16	12.5	710	14	CB45570	696934 NA
C 377	16	12.5	609	14	CY786085	sat43Q03.	CY786085	sat43Q03.	450	16	12.5	713	28	BH956168	odi51e07.
C 378	16	12.5	610	14	C79015	Mous	C79015	C79015	C 451	16	12.5	717	28	BH488963	BOHN253TR
C 379	16	12.5	610	14	C79015	Mous	C79015	C79015	C 452	16	12.5	719	10	BG574022	602595867
C 380	16	12.5	612	9	AI980497	pat_pk002	AI980497	pat_pk002	453	16	12.5	719	14	CA919548	EST637266
C 381	16	12.5	614	28	BH098299	RPCI-24-3	BH098299	RPCI-24-3	C 454	16	12.5	720	28	BH959652	oddsIn09.
C 382	16	12.5	615	12	BI163230	RE02574.5	BI163230	RE02574.5	455	16	12.5	721	28	AZ326150	IM0048G10
C 383	16	12.5	616	9	AV743206	AV743206	AV743206	AV743206	C 456	16	12.5	722	12	BI088772	BI088772
C 384	16	12.5	617	28	AO692234	HS 5407.A	AO692234	HS 5407.A	457	16	12.5	725	14	CB498877	CB498877
C 385	16	12.5	621	12	BJ081652	BJ081652	BJ081652	BJ081652	C 458	16	12.5	726	28	BH692334	BH692334
C 386	16	12.5	621	28	BH357268	BH357268	BH357268	BH357268	C 459	16	12.5	727	28	BH982888	BH982888
C 387	16	12.5	622	10	BF297771	056PBA09	BF297771	056PBA09	460	16	12.5	729	29	BX245638	Danio rer
C 388	16	12.5	624	12	BM975905	UI-CF-EN1	BM975905	UI-CF-EN1	461	16	12.5	733	28	BH460319	BOGZ929TR
C 389	16	12.5	628	14	CB444527	695698 MA	CB444527	695698 MA	C 462	16	12.5	734	28	BH663528	BOHV124TF
C 390	16	12.5	628	14	CB444527	695698 MA	CB444527	695698 MA	463	16	12.5	734	28	BH663528	BOHV124TF
C 391	16	12.5	629	13	CB577813	AMGNNUC:N	CB577813	AMGNNUC:N	464	16	12.5	735	29	CNS04K3D	AL294322
C 392	16	12.5	629	13	CB524786	NISC no07	CB524786	NISC no07	C 465	16	12.5	735	13	BQ516329	EST63644
C 393	16	12.5	631	28	BQ581912	il110f07.Y	BQ581912	il110f07.Y	C 466	16	12.5	735	28	BH459620	BOSTE29TR
C 394	16	12.5	632	28	BZ020572	oe90c06.	BZ020572	oe90c06.	C 467	16	12.5	736	14	CB459909	BOHS282TR
C 395	16	12.5	635	28	AQ964176	LEGR71TF	AQ964176	LEGR71TF	C 468	16	12.5	738	28	BH607695	BOQJL26TR
C 396	16	12.5	635	29	BX174590	Danio rer	BX174590	Danio rer	469	16	12.5	738	13	BH462802	603

C 516	16	12.5	815	28	BH646077	BH646077 BOHXD34TR	589	29	CC229889	CC229889 CH261-46M
C 517	16	12.5	816	13	BH747764	BH747764 CH3#015_D	C 590	29	CC217302	CC217302 CH261-69C
C 518	16	12.5	816	29	CA339617	CA339617 PUHJD91TD	C 591	14	CB984237	CB984237 AGENCOURT
C 519	16	12.5	820	12	B1866777	B1866777 603391187	C 592	29	CC235303	CC235303 CH261-71P
C 520	16	12.5	820	28	BH4077152	BH4077152 BOGZRO8TR	C 593	12	BM556256	BM556256 AGENCOURT
C 521	16	12.5	821	13	BH4077679	BH4077679 EX077679	C 594	29	CC256316	CC256316 CH261-166
C 522	16	12.5	826	28	BH2143603	BH2143603 CH230-319	C 595	10	BG393381	BG393381 602411724
C 523	16	12.5	830	28	BH478244	BH478244 BOGHE69TR	C 596	9	AL563446	AL563446 AL563446
C 524	16	12.5	831	13	BH747765	BH747765 CH3#015_D	C 597	13	EX447720	EX447720 EX447720
C 525	16	12.5	834	14	CD246007	CD246007 AGENCOURT	C 598	29	CC231094	CC231094 CH261-57G
C 526	16	12.5	837	28	AL716352	AL716352 SP 0141_A	C 599	29	CC221199	CC221199 CH261-183
C 527	16	12.5	839	28	BH484202	BH484202 BOHQ182FT	C 600	12	BM549575	BM549575 AGENCOURT
C 528	16	12.5	843	28	BH2144629	BH2144629 CH230-246	C 601	16	AK030666	AK030666 Mus muscu
C 529	16	12.5	847	28	AO748651	AO748651 HS 5573_A	C 602	16	AK029065	AK029065 Mus muscu
C 530	16	12.5	850	12	BT554161	BT554161 603235584	C 603	16	AK048716	AK048716 Mus muscu
C 531	16	12.5	852	12	BH686167	BH686167 602789048	C 604	15	70 28	AZ918371 1006004E0
C 532	16	12.5	852	13	BU718007	BU718007 SJM2HUF0T	C 605	15	112 14	T87201 yd91h02.s1
C 533	16	12.5	853	29	CC385616	CC385616 PUHPR48TD	C 606	15	114 10	BE685469 187721_MA
C 534	16	12.5	854	14	CB314141	CB314141 AGENCOURT	C 607	15	117 121	AZ918668 1006004H1
C 535	16	12.5	855	29	CC339641	CC339641 OGIAF36TH	C 608	29	CNS000MD	AL080731 Arabidops
C 536	16	12.5	856	28	BH154903	BH154903 ENTOW55TF	C 609	15	134 29	AL770141 Arabidops
C 537	16	12.5	856	28	BH154903	BH154903 ENTOW55TF	C 610	15	140 10	BG549032 947072D09
C 538	16	12.5	868	28	BH656730	BH656730 BOHWA24TF	C 611	142 12	BI005004	BI005004 RC4-HN017
C 539	16	12.5	869	29	CF422457	CF422457 PUHPR48TD	C 612	147 10	BE478290	BE478290 162473_BA
C 540	16	12.5	881	10	BF666305	BF666305 602119223	C 613	153 28	AQ046820	AQ046820 RFC111-33
C 541	16	12.5	882	10	BF237259	BF237259 602025420	C 614	155 9	AI874246	AI874246 tz63d04.x
C 542	16	12.5	882	13	BU504885	BU504885 AGENCOURT	C 615	157 28	AZ111973	AZ111973 RPCI-23-9
C 543	16	12.5	889	13	BX350080	BX350080 BX350080	C 616	159 10	BM052174	BM052174 BE052174
C 544	16	12.5	890	28	AZ538563	AZ538563 ENTGF68TR	C 617	166 12	BM084268	BM084268 K-ES10095
C 545	16	12.5	892	12	B1729972	B1729972 603350065	C 618	169 13	BY014325	BY014325 BY014325
C 546	16	12.5	894	10	BG567446	BG567446 602585969	C 619	172 28	AZ880099	AZ880099 RPCI-23-2
C 547	16	12.5	900	28	BH163556	BH163556 ENTRO35TF	C 620	175 9	AI601180	AI601180 ar89f11.x
C 548	16	12.5	903	28	BH160604	BH160604 ENTRO94TR	C 621	176 14	Z36327	Z36327 MW234_MVM (
C 549	16	12.5	907	13	BU508425	BU508425 AGENCOURT	C 622	180 14	CB550403	CB550403 MWPL0001
C 550	16	12.5	914	14	CA489412	CA489412 AGENCOURT	C 623	183 13	BU334974	BU334974 603803984
C 551	16	12.5	920	28	AZ544304	AZ544304 ENTFC31TF	C 624	184 10	BE081197	BE081197 QVI-ET063
C 552	16	12.5	922	9	AL527447	AL527447 AL527447	C 625	185 28	AO637407	AO637407 RPCI-11-4
C 553	16	12.5	927	29	CC134385	CC134385 NDL_60D18	C 626	198 10	BE986406	BE986406 UI-M-CG0P
C 554	16	12.5	934	13	BX326616	BX326616 BX326616	C 627	201 12	BI007543	BI007543 MR1-RT007
C 555	16	12.5	935	10	BG394315	BG394315 602456525	C 628	211 10	BF447609	BF447609 7G93B06.x
C 556	16	12.5	948	12	B1408725	B1408725 602945668	C 629	216 9	AA974169	AA974169 oq13f03.s
C 557	16	12.5	950	13	BU153444	BU153444 AGENCOURT	C 630	216 10	BS010806	BS010806 BS010806
C 558	16	12.5	954	28	AZ686121	AZ686121 ENTFC26TR	C 631	216 10	BB466378	BB466378 BB466378
C 559	16	12.5	954	29	CNS037H8	AL231317 Tetradon	C 632	226 10	BB249522	BB249522 BB249522
C 560	16	12.5	957	10	BE795646	BE795646 601590629	C 633	227 9	AV141631	AV141631 AV141631
C 561	16	12.5	960	10	BH970186	BH970186 602734404	C 634	227 13	BQ369276	BQ369276 PM2-GN051
C 562	16	12.5	974	13	BQ364411	BQ364411 AGENCOURT	C 635	231 14	R43148	R43148 y91h0h03.s1
C 563	16	12.5	975	10	BF180931	BF180931 601807171	C 636	233 28	AZ654984	AZ654984 IM0529112
C 564	16	12.5	983	13	BX385682	BX385682 BX385682	C 637	239 9	AI789970	AI789970 ue63e01.r
C 565	16	12.5	984	13	BU137496	BU137496 603124860	C 638	241 10	BB170789	BB170789 BB170789
C 566	16	12.5	985	11	CNS09B2C	BX050416 Single re	C 639	245 9	AI037645	AI037645 uh22g11.r
C 567	16	12.5	986	13	BO711807	BO711807 AGENCOURT	C 640	245 10	BE529972	BE529972 M76121STM
C 568	16	12.5	988	14	CD253432	CD253432 AGENCOURT	C 641	245 10	BE529972	BE529972 M76121STM
C 569	16	12.5	993	29	CC189193	CC189193 CH261-35P	C 642	247 9	AA287044	AA287044 zss7e09.s
C 570	16	12.5	994	13	BU050965	BU050965 AGENCOURT	C 643	247 28	AZ256809	AZ256809 RPCI-23-1
C 571	16	12.5	1002	12	BM562117	BM562117 AGENCOURT	C 644	248 10	BE530158	BE530158 M76F01STM
C 572	16	12.5	1009	12	BM914450	BM914450 AGENCOURT	C 645	250 9	AW478738	AW478738 21145_MAR
C 573	16	12.5	1032	12	BM557159	BM557159 AGENCOURT	C 646	251 9	AV356079	AV356079 AV356079
C 574	16	12.5	1035	12	BM914584	BM914584 AGENCOURT	C 647	253 28	AZ700689	AZ700689 RPCI-23-2
C 575	16	12.5	1040	29	CC238797	CC238797 CH261-120	C 648	254 9	AM478739	AM478739 21150_MAR
C 576	16	12.5	1042	29	CC190686	CC190686 CH261-65B	C 649	255 14	CB483116	CB483116 jns03_D07
C 577	16	12.5	1049	12	BM915136	BM915136 AGENCOURT	C 650	256 14	R85570	R85570 yt67a05.s1
C 578	16	12.5	1051	10	BG259445	BG259445 602378541	C 651	259 9	AW337098	AW337098 22199_MAR
C 579	16	12.5	1053	13	BU723277	BU723277 SJMA7E03	C 652	260 9	AV334278	AV334278 AV334278
C 580	16	12.5	1060	12	BM468562	BM468562 AGENCOURT	C 653	262 14	CA347500	CA347500 678641_NC
C 581	16	12.5	1061	13	BQ067926	BQ067926 AGENCOURT	C 654	264 10	BF457418	BF457418 UI-M-BZ1-
C 582	16	12.5	1065	13	BX390990	BX390990 BX390990	C 655	265 10	AW876125	AW876125 PM4-PT001
C 583	16	12.5	1074	29	CNS05LXU	AL343371 Tetradon	C 656	267 10	BF741682	BF741682 CM4-HB002
C 584	16	12.5	1083	12	BM553502	BM553502 AGENCOURT	C 657	268 9	AW166742	AW166742 xg67g05.x
C 585	16	12.5	1088	29	CC246519	CC246519 CH261-87H	C 658	268 28	BH297072	BH297072 CH230-143
C 586	16	12.5	1108	12	BM460594	BM460594 AGENCOURT	C 659	269 10	BB419722	BB419722 CH230-143
C 587	16	12.5	1110	29	CNS07E98	AL41474 T3 end of	C 660	274 10	BE046965	BE046965 BE046965
C 588	16	12.5	1113	12	BM553216	BM553216 AGENCOURT	C 661	275 9	AI502054	AI502054 UI-R-C1-j



C 662	15	11.7	275	9	AU258890	AU258890	BH012298	TDGAH63TH
C 663	15	11.7	277	28	BH452885	BH452885	BH012298	TDGAH63TH
C 664	15	11.7	280	10	BE205966	BE205966	BH012298	TDGAH63TH
C 665	15	11.7	281	14	CB130557	CB130557	BH012298	TDGAH63TH
C 666	15	11.7	282	28	AQ236220	AQ236220	BH012298	TDGAH63TH
C 667	15	11.7	284	10	BE756276	BE756276	BH012298	TDGAH63TH
C 668	15	11.7	284	29	CC199215	CC199215	BH012298	TDGAH63TH
C 669	15	11.7	285	10	BE155656	BE155656	BH012298	TDGAH63TH
C 670	15	11.7	287	28	AQ066491	AQ066491	BH012298	TDGAH63TH
C 671	15	11.7	288	10	BF323314	BF323314	BH012298	TDGAH63TH
C 672	15	11.7	288	10	BF323314	BF323314	BH012298	TDGAH63TH
C 673	15	11.7	291	29	AX002632	AX002632	BH012298	TDGAH63TH
C 674	15	11.7	293	9	AW319862	AW319862	BH012298	TDGAH63TH
C 675	15	11.7	293	13	BQ350939	BQ350939	BH012298	TDGAH63TH
C 676	15	11.7	294	10	BE005319	BE005319	BH012298	TDGAH63TH
C 677	15	11.7	294	28	AZ971453	AZ971453	BH012298	TDGAH63TH
C 678	15	11.7	295	14	CB473461	CB473461	BH012298	TDGAH63TH
C 679	15	11.7	296	9	AW652591	AW652591	BH012298	TDGAH63TH
C 680	15	11.7	296	14	N76236	N76236	BH012298	TDGAH63TH
C 681	15	11.7	297	9	AW347727	AW347727	BH012298	TDGAH63TH
C 682	15	11.7	298	10	BE091466	BE091466	BH012298	TDGAH63TH
C 683	15	11.7	298	10	BE542880	BE542880	BH012298	TDGAH63TH
C 684	15	11.7	298	28	AQ094712	AQ094712	BH012298	TDGAH63TH
C 685	15	11.7	300	10	BE246811	BE246811	BH012298	TDGAH63TH
C 686	15	11.7	302	28	BH081261	BH081261	BH012298	TDGAH63TH
C 687	15	11.7	304	10	BF449462	BF449462	BH012298	TDGAH63TH
C 688	15	11.7	304	28	AZ147075	AZ147075	BH012298	TDGAH63TH
C 689	15	11.7	306	10	BF361385	BF361385	BH012298	TDGAH63TH
C 690	15	11.7	307	10	BE491533	BE491533	BH012298	TDGAH63TH
C 691	15	11.7	308	10	BF927534	BF927534	BH012298	TDGAH63TH
C 692	15	11.7	309	10	BE458252	BE458252	BH012298	TDGAH63TH
C 693	15	11.7	309	14	T52431	T52431	BH012298	TDGAH63TH
C 694	15	11.7	310	9	AW050528	AW050528	BH012298	TDGAH63TH
C 695	15	11.7	311	14	CD284714	CD284714	BH012298	TDGAH63TH
C 696	15	11.7	314	10	BE221353	BE221353	BH012298	TDGAH63TH
C 697	15	11.7	315	12	BI031848	BI031848	BH012298	TDGAH63TH
C 698	15	11.7	315	13	BY329141	BY329141	BH012298	TDGAH63TH
C 699	15	11.7	316	28	AZ153684	AZ153684	BH012298	TDGAH63TH
C 700	15	11.7	317	10	BE446175	BE446175	BH012298	TDGAH63TH
C 701	15	11.7	319	12	BE289808	BE289808	BH012298	TDGAH63TH
C 702	15	11.7	319	14	Z41246	Z41246	BH012298	TDGAH63TH
C 703	15	11.7	320	10	BE728009	BE728009	BH012298	TDGAH63TH
C 704	15	11.7	322	10	BA434144	BA434144	BH012298	TDGAH63TH
C 705	15	11.7	322	10	BA462613	BA462613	BH012298	TDGAH63TH
C 706	15	11.7	322	10	BE011480	BE011480	BH012298	TDGAH63TH
C 707	15	11.7	322	13	BY012217	BY012217	BH012298	TDGAH63TH
C 708	15	11.7	322	28	AQ094327	AQ094327	BH012298	TDGAH63TH
C 709	15	11.7	324	9	AW331825	AW331825	BH012298	TDGAH63TH
C 710	15	11.7	324	9	AW455828	AW455828	BH012298	TDGAH63TH
C 711	15	11.7	324	10	BH388337	BH388337	BH012298	TDGAH63TH
C 712	15	11.7	324	10	BA490691	BA490691	BH012298	TDGAH63TH
C 713	15	11.7	325	29	EX002633	EX002633	BH012298	TDGAH63TH
C 714	15	11.7	326	10	BF679508	BF679508	BH012298	TDGAH63TH
C 715	15	11.7	327	9	AV671176	AV671176	BH012298	TDGAH63TH
C 716	15	11.7	327	13	BY329641	BY329641	BH012298	TDGAH63TH
C 717	15	11.7	329	10	BE654367	BE654367	BH012298	TDGAH63TH
C 718	15	11.7	333	9	AA142324	AA142324	BH012298	TDGAH63TH
C 719	15	11.7	333	10	BG379900	BG379900	BH012298	TDGAH63TH
C 720	15	11.7	333	10	BA497818	BA497818	BH012298	TDGAH63TH
C 721	15	11.7	333	28	BH073850	BH073850	BH012298	TDGAH63TH
C 722	15	11.7	335	9	AW324385	AW324385	BH012298	TDGAH63TH
C 723	15	11.7	335	14	CB161095	CB161095	BH012298	TDGAH63TH
C 724	15	11.7	336	10	BF989885	BF989885	BH012298	TDGAH63TH
C 725	15	11.7	336	10	BE611265	BE611265	BH012298	TDGAH63TH
C 726	15	11.7	336	14	CB296295	CB296295	BH012298	TDGAH63TH
C 727	15	11.7	337	13	BY150890	BY150890	BH012298	TDGAH63TH
C 728	15	11.7	338	10	BF989882	BF989882	BH012298	TDGAH63TH
C 729	15	11.7	338	13	BY010662	BY010662	BH012298	TDGAH63TH
C 730	15	11.7	338	14	C95449	C95449	BH012298	TDGAH63TH
C 731	15	11.7	341	9	AV219879	AV219879	BH012298	TDGAH63TH
C 732	15	11.7	342	10	BF710549	BF710549	BH012298	TDGAH63TH
C 733	15	11.7	342	28	AZ318026	AZ318026	BH012298	TDGAH63TH
C 734	15	11.7	344	9	AW652586	AW652586	BH012298	TDGAH63TH
C 735	15	11.7	344	28	AZ605695	AZ605695	BH012298	TDGAH63TH
C 736	15	11.7	345	12	BG997011	BG997011	BH012298	TDGAH63TH
C 737	15	11.7	345	13	BY560238	BY560238	BH012298	TDGAH63TH
C 738	15	11.7	346	10	BF170589	BF170589	BH012298	TDGAH63TH
C 739	15	11.7	346	12	BI029243	BI029243	BH012298	TDGAH63TH
C 740	15	11.7	346	28	AQ073039	AQ073039	BH012298	TDGAH63TH
C 741	15	11.7	347	10	BE044991	BE044991	BH012298	TDGAH63TH
C 742	15	11.7	347	14	CA783092	CA783092	BH012298	TDGAH63TH
C 743	15	11.7	348	9	AI504118	AI504118	BH012298	TDGAH63TH
C 744	15	11.7	349	13	BQ330083	BQ330083	BH012298	TDGAH63TH
C 745	15	11.7	349	28	AQ268477	AQ268477	BH012298	TDGAH63TH
C 746	15	11.7	350	9	AU251608	AU251608	BH012298	TDGAH63TH
C 747	15	11.7	350	13	BY498258	BY498258	BH012298	TDGAH63TH
C 748	15	11.7	351	9	AA138424	AA138424	BH012298	TDGAH63TH
C 749	15	11.7	353	9	AW431609	AW431609	BH012298	TDGAH63TH
C 750	15	11.7	353	28	AZ498558	AZ498558	BH012298	TDGAH63TH
C 751	15	11.7	353	29	CNS01M30	CNS01M30	BH012298	TDGAH63TH
C 752	15	11.7	355	9	AA250352	AA250352	BH012298	TDGAH63TH
C 753	15	11.7	355	10	BF768968	BF768968	BH012298	TDGAH63TH
C 754	15	11.7	355	13	BY293943	BY293943	BH012298	TDGAH63TH
C 755	15	11.7	355	14	F081174	F081174	BH012298	TDGAH63TH
C 756	15	11.7	356	9	AV697338	AV697338	BH012298	TDGAH63TH
C 757	15	11.7	359	13	BQ351080	BQ351080	BH012298	TDGAH63TH
C 758	15	11.7	360	10	BF469595	BF469595	BH012298	TDGAH63TH
C 759	15	11.7	360	14	HS7075	HS7075	BH012298	TDGAH63TH
C 760	15	11.7	361	13	BQ491298	BQ491298	BH012298	TDGAH63TH
C 761	15	11.7	362	28	AZ429371	AZ429371	BH012298	TDGAH63TH
C 762	15	11.7	362	28	AQ333280	AQ333280	BH012298	TDGAH63TH
C 763	15	11.7	364	9	AW010125	AW010125	BH012298	TDGAH63TH
C 764	15	11.7	366	28	BH680870	BH680870	BH012298	TDGAH63TH
C 765	15	11.7	367	13	BY294391	BY294391	BH012298	TDGAH63TH
C 766	15	11.7	368	14	CD541095	CD541095	BH012298	TDGAH63TH
C 767	15	11.7	369	9	AA036003	AA036003	BH012298	TDGAH63TH
C 768	15	11.7	369	14	CB923372	CB923372	BH012298	TDGAH63TH
C 769	15	11.7	370	12	BI011318	BI011318	BH012298	TDGAH63TH
C 770	15	11.7	370	28	BH072703	BH072703	BH012298	TDGAH63TH
C 771	15	11.7	371	10	BF146918	BF146918	BH012298	TDGAH63TH
C 772	15	11.7	372	13	BY230457	BY230457	BH012298	TDGAH63TH
C 773	15	11.7	372	14	CB479190	CB479190	BH012298	TDGAH63TH
C 774	15	11.7	373	13	BY428633	BY428633	BH012298	TDGAH63TH
C 775	15	11.7	374	12	BM121411	BM121411	BH012298	TDGAH63TH
C 776	15	11.7	374	13	BY414152	BY414152	BH012298	TDGAH63TH
C 777	15	11.7	374	13	BY440987	BY440987	BH012298	TDGAH63TH
C 778	15	11.7	374	29	CC076785	CC076785	BH012298	TDGAH63TH
C 779	15	11.7	375	10	BE045120	BE045120	BH012298	TDGAH63TH
C 780	15	11.7	375	14	CB691967	CB691967	BH012298	TDGAH63TH
C 781	15	11.7	375	14	CB975357	CB975357	BH012298	TDGAH63TH
C 782	15	11.7	376	28	AQ084624	AQ084624	BH012298	TDGAH63TH
C 783	15	11.7	376	28	AQ084624	AQ084624	BH012298	TDGAH63TH
C 784	15	11.7	376	29	BZ871864	BZ871864	BH012298	TDGAH63TH
C 785	15	11.7	376	29	CNS00P23	CNS00P23	BH012298	TDGAH63TH
C 786	15	11.7	377	10	BE791718	BE791718	BH012298	TDGAH63TH
C 787	15	11.7	378	13	BY375063	BY375063	BH012298	TDGAH63TH
C 788	15	11.7	378	14	CB474390	CB474390	BH012298	TDGAH63TH
C 789	15	11.7	378	14	CB479246	CB479246	BH012298	TDGAH63TH
C 790	15	11.7	380	9	AA214228	AA214228	BH012298	TDGAH63TH
C 791	15	11.7	380	14	T97394	T97394	BH012298	TDGAH63TH
C 792	15	11.7	381	13	BY540933	BY540933	BH012298	TDGAH63TH
C 793	15	11.7	382	10	BG040917	BG040917	BH012298	TDGAH63TH
C 794	15	11.7	382	12	BI050362	BI050362	BH012298	TDGAH63TH
C 795	15	11.7	382	14	CB475716	CB475716	BH012298	TDGAH63TH
C 796	15	11.7	382	14	CB475716	CB475716	BH012298	TDGAH6

808	15	11.7	386	13	BY528383	BY528383	881	15	11.7	417	13	BY622004	BY622004
c 809	15	11.7	386	13	BY563021	BY563021	882	15	11.7	417	13	BY624338	BY624338
810	15	11.7	386	29	FR0002150	Z86533	c 883	15	11.7	417	28	AZ080858	RPCI-23-4
811	15	11.7	387	13	BY659385	BY659385	c 884	15	11.7	417	28	AZ465931	1M0276N10
812	15	11.7	387	28	AZ345072	AZ345072	c 885	15	11.7	417	28	AQ380056	RPCI11-16
c 813	15	11.7	388	9	AA175574	ms33d03.r	c 886	15	11.7	418	9	AA692616	AA692616
c 814	15	11.7	388	10	AA898772	CM0-NN007	c 887	15	11.7	418	9	AA813492	AA813492
c 815	15	11.7	389	12	BP015616	BP015616	c 888	15	11.7	418	12	BM751334	BM751334
816	15	11.7	390	10	BF742038	CM4-HB002	c 889	15	11.7	419	9	AI020240	ub26d01.r
817	15	11.7	390	12	B244986	AR05F05L	c 890	15	11.7	419	13	BY109282	BY109282
c 818	15	11.7	390	13	BQ350938	OVI-HT064	c 891	15	11.7	420	14	T38099	EST103460 S
c 819	15	11.7	390	13	BY628039	BY628039	c 892	15	11.7	421	29	CNS000A9	AL082959 Arabidops
c 820	15	11.7	390	14	CA960911	CA960911	c 893	15	11.7	422	13	BY240655	BY240655
c 821	15	11.7	393	10	BF524856	BF524856	c 894	15	11.7	423	13	BY296283	BY296283
c 822	15	11.7	393	10	BE044983	hm25c10.x	c 895	15	11.7	423	13	BY516789	BY516789
c 823	15	11.7	393	13	BQ368299	PM3-GN051	c 896	15	11.7	423	28	AZ561246	RPCI-23-2
c 824	15	11.7	393	13	BU428608	UI-HP-ENO	c 897	15	11.7	424	9	AA624354	AA624354
c 825	15	11.7	393	28	AO609908	HS-5089.A	c 898	15	11.7	424	10	BE352329	BE352329
c 826	15	11.7	394	9	AA087037	mk20d05.r	c 899	15	11.7	424	10	BE576668	BE576668
c 827	15	11.7	394	13	BY058715	BY058715	c 900	15	11.7	424	12	BM956435	BM956435
828	15	11.7	394	28	BY748607	BY748607	c 901	15	11.7	424	13	BY266551	BY266551
829	15	11.7	395	13	BY440561	BY440561	c 902	15	11.7	424	13	BY363794	BY363794
830	15	11.7	395	14	W63738	zd30c08.s1	c 903	15	11.7	425	13	BY276028	BY276028
831	15	11.7	396	10	BE931968	BE931968	c 904	15	11.7	425	14	H00599	YJ25c04.s1
832	15	11.7	396	10	BE794302	BE794302	c 905	15	11.7	425	28	AZ070155	RPCI-23-3
c 833	15	11.7	396	28	BE188745	BE188745	c 906	15	11.7	426	9	AI469881	AI469881
c 834	15	11.7	396	29	CNS07SOY	AL625684	c 907	15	11.7	426	9	AV878810	AV878810
c 835	15	11.7	397	14	C84576	C84576	c 908	15	11.7	427	12	BM794316	BM794316
836	15	11.7	398	9	AI143085	qb71h05.x	c 909	15	11.7	427	12	BM794316	K-ESN0075
c 837	15	11.7	398	14	CE071188	CE071188	c 910	15	11.7	428	9	AA178131	mt14b12.z
c 838	15	11.7	398	14	R76386	R76386	c 911	15	11.7	429	10	BB864613	BB864613
839	15	11.7	399	9	AM620239	si93e10.y	c 912	15	11.7	429	12	CA283931	CA283931
840	15	11.7	399	28	AQ219475	AQ219475	c 913	15	11.7	429	14	CA883318	CA883318
841	15	11.7	400	9	AI003185	ai21h04.s	c 914	15	11.7	430	10	BF544351	BF544351
c 842	15	11.7	400	9	AI045761	UI-R-CL-k	c 915	15	11.7	431	12	BI133674	BI133674
c 843	15	11.7	400	9	AE293241	UI-H-BI2-	c 916	15	11.7	431	13	BY296833	BY296833
c 844	15	11.7	400	13	BY494112	BY494112	c 917	15	11.7	431	28	AQ045723	AQ045723
c 845	15	11.7	401	10	BF7464319	BF7464319	c 918	15	11.7	431	28	AQ060623	AQ060623
846	15	11.7	401	10	BE739430	BE739430	c 919	15	11.7	431	28	AQ774173	AQ774173
847	15	11.7	401	13	BQ704983	Bn01_02a2	c 920	15	11.7	432	12	BI563329	BI563329
c 848	15	11.7	401	14	CA961446	CA961446	c 921	15	11.7	432	12	CA870413	CA870413
c 849	15	11.7	402	10	BF377871	BF377871	c 922	15	11.7	432	28	BH219660	BH219660
c 850	15	11.7	402	10	BE731927	BE731927	c 923	15	11.7	433	13	EU721052	EU721052
c 851	15	11.7	402	13	BY524063	BY524063	c 924	15	11.7	433	9	AI802003	AI802003
c 852	15	11.7	403	13	BY533230	BY533230	c 925	15	11.7	434	10	BB787978	BB787978
c 853	15	11.7	403	9	AA465621	AA465621	c 926	15	11.7	434	13	BY494462	BY494462
c 854	15	11.7	405	12	BM524795	BM524795	c 927	15	11.7	434	28	B76751	B76751
c 855	15	11.7	405	13	BY485465	BY485465	c 928	15	11.7	434	28	BZ154547	BZ154547
c 856	15	11.7	405	28	BE348303	BE348303	c 929	15	11.7	434	28	AQ498878	AQ498878
c 857	15	11.7	406	14	CA959915	CA959915	c 930	15	11.7	435	9	AA059765	AA059765
c 858	15	11.7	407	12	BM878268	BM878268	c 931	15	11.7	435	9	AW699429	AW699429
c 859	15	11.7	408	9	AV748127	AV748127	c 932	15	11.7	435	10	BB700062	BB700062
c 860	15	11.7	409	13	BY549508	BY549508	c 933	15	11.7	436	12	EG814532	EG814532
c 861	15	11.7	410	14	CA541045	CA541045	c 934	15	11.7	437	9	AA773716	AA773716
c 862	15	11.7	411	13	BY511801	BY511801	c 935	15	11.7	437	9	AA895151	AA895151
863	15	11.7	411	13	BY652472	BY652472	c 936	15	11.7	437	14	CA597232	CA597232
864	15	11.7	411	29	CO054646	CO054646	c 937	15	11.7	437	14	W59109	W59109
865	15	11.7	412	13	BY495195	BY495195	c 938	15	11.7	437	28	AQ612443	AQ612443
c 866	15	11.7	412	13	BY514474	BY514474	c 939	15	11.7	438	9	AU014527	AU014527
c 867	15	11.7	413	13	EQ350941	EQ350941	c 940	15	11.7	438	10	AW911992	AW911992
c 868	15	11.7	413	13	BY289370	BY289370	c 941	15	11.7	438	13	BY449990	BY449990
869	15	11.7	413	13	BY516297	BY516297	c 942	15	11.7	438	13	BY500967	BY500967
870	15	11.7	414	9	AV801484	AV801484	c 943	15	11.7	438	13	BY537200	BY537200
871	15	11.7	414	10	BG725743	BG725743	c 944	15	11.7	438	28	AZ905100	AZ905100
c 872	15	11.7	415	9	AA276947	AA276947	c 945	15	11.7	438	28	AQ616808	AQ616808
c 873	15	11.7	415	10	AM876179	AM876179	c 946	15	11.7	439	9	AA420076	AA420076
874	15	11.7	415	10	BE829048	BE829048	c 947	15	11.7	439	10	BE209435	BE209435
875	15	11.7	416	10	BE059001	BE059001	c 948	15	11.7	439	14	CA567851	CA567851
c 876	15	11.7	416	13	BY284684	BY284684	c 949	15	11.7	440	28	BH033196	BH033196
c 877	15	11.7	416	13	BY492196	BY492196	c 950	15	11.7	440	28	BH033196	BH033196
878	15	11.7	416	13	BY511700	BY511700	c 951	15	11.7	440	29	CC156238	CC156238
c 879	15	11.7	417	9	AI840621	AI840621	c 952	15	11.7	441	10	BB819491	BB819491
c 880	15	11.7	417	12	BG960699	BG960699	c 953	15	11.7	441	10	BF147282	BF147282

```

c 954      15 11.7 442 14  CR885278
c 955      15 11.7 442 28  A0863979
c 956      15 11.7 442 28  B2193189
c 957      15 11.7 443 10  A0876198
c 958      15 11.7 443 28  A0560224
c 959      15 11.7 444 28  A0535027
c 960      15 11.7 444 29  C050788V
c 961      15 11.7 445 9  A0682543
c 962      15 11.7 445 10  B0748575
c 963      15 11.7 445 12  BG797417
c 964      15 11.7 445 28  A2167623
c 965      15 11.7 446 9  AA773807
c 966      15 11.7 446 10  BF437030
c 967      15 11.7 446 13  BY460136
c 968      15 11.7 447 9  AL919527
c 969      15 11.7 447 9  AW653125
c 970      15 11.7 447 13  BU207784
c 971      15 11.7 447 14  CA562883
c 972      15 11.7 448 28  A0511309
c 973      15 11.7 448 28  A2170554
c 974      15 11.7 448 29  B2901440
c 975      15 11.7 449 9  AV636587
c 976      15 11.7 449 13  BY242508
c 977      15 11.7 449 14  CA872501
c 978      15 11.7 449 14  CA876121
c 979      15 11.7 449 14  CD541700
c 980      15 11.7 450 9  AW576588
c 981      15 11.7 450 28  A0984454
c 982      15 11.7 450 28  B2173452
c 983      15 11.7 451 10  BF796106
c 984      15 11.7 451 28  A0402980
c 985      15 11.7 452 9  AA794748
c 986      15 11.7 452 9  AI316834
c 987      15 11.7 452 9  AL793883
c 988      15 11.7 452 10  BB831714
c 989      15 11.7 452 13  BY555365
c 990      15 11.7 453 9  AA682616
c 991      15 11.7 453 9  AA755806
c 992      15 11.7 453 9  AI375009
c 993      15 11.7 453 10  BB826739
c 994      15 11.7 453 13  BY239467
c 995      15 11.7 453 14  CA541085
c 996      15 11.7 454 10  BF021806
c 997      15 11.7 454 14  W59112
c 998      15 11.7 455 9  AI952839
c 999      15 11.7 455 10  BB839121
c 1000     15 11.7 455 14  CB741760

```

## ALIGNMENTS

```

RESULT 1
AUI00197      300 bp mRNA linear EST 05-APR-2001
LOCUS      AUI00197 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION  RECO1067 similar to Homo sapiens meglin mRNA, mRNA sequence.
ACCESSION  AUI00197
VERSION    AUI00197.1 GI:13551326
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 300)
AUTHORS     Suzuki.Y., Tsunoda.T., Taira.H., Mizushima-Sugano.J., Sese.J., Hata
            .H., Oka.T., Isogai.T., Tanaka.T., Nakamura.Y., Morishita.S., Okubo
            .K., Suyama.A. and Sugano.S.
TITLE       In silico mapping of the 5'-ends of human mRNAs using full-length
            enriched and 5'-end enriched cDNA libraries constructed by
            oligo-capping method
JOURNAL     Unpublished
COMMENT     Contact: Yutaka Suzuki

```

```

Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp
Suzuki.Y., Yoshitomo-Nakagawa.K., Maruyama.K., Suyama.A. and Sugano
.S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             Location/Qualifiers
     source            1..300
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="REC01067"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      88 a 63 c 70 g 79 t
ORIGIN
Query Match      23.4%; Score 30; DB 9; Length 300;
Best Local Similarity 100.0%; Pred.No.1.5e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 GAGTCATAGGAGGAGCCATCCCGAAGCCAG 128
      |||
Db 1 GAGTCATAGGAGGAGCCATCCCGAAGCCAG 30

RESULT 2
AA975162      142 bp mRNA linear EST 22-MAY-1998
LOCUS      AA975162 om99e07.61 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1555332 3',
DEFINITION  mRNA sequence.
ACCESSION  AA975162
VERSION    AA975162.1 GI:3150954
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 142)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LINL at:
            www.bio.lnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 53.
Location/Qualifiers
     source            1..142
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1555332"
                     /lab_host="DH10B"
                     /clone_lib="NCI CGAP Kid3"
                     /note="Organ: kidney; Vector: pRT3D-Pac (Pharmacia) with
            a modified polylinker; Site:1: Not 1; Site:2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer,
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not
            I and Eco RI sites of the modified pRT3 vector. mRNA
            source: 2 pooled kidneys. Library went through one round
            of normalization. Library constructed by Bento Soares and
            M. Fatima Bonaldo. "
BASE COUNT      65 a 27 c 13 g 37 t

```

ORIGIN		Query Match		14.8%; Score 19; DB 9; Length 142;		Best Local Similarity 100.0%; Pred. No. 11;		Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	53	AAACCTAAATGCTTATAA 71							
Db	42	AAACCTAAATGCTTATAA 60							
RESULT 3		BF365378/c		294 bp		mrna		linear	
LOCUS		QV4-NT0028-080700-286-c08 NT0028 Homo sapiens cDNA, mRNA sequence.		EST 24-NOV-2000					
DEFINITION		BF365378							
ACCESSION		BF365378.1		GI:11327403					
VERSION		EST.							
KEYWORDS		Homo sapiens (human)							
SOURCE		Homo sapiens							
ORGANISM		Homo sapiens							
REFERENCE		Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.							
AUTHORS		1 (bases 1 to 294)							
TITLE		Shotgun sequencing of the human transcriptome with ORF expressed							
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)							
MEDLINE		20202663							
PUBMED		10737800							
COMMENT		Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-NT0028-080700-286-c08&t3=2000-07-08&t4=1) Seq primer: puc 18 forward High quality sequence start: 16 High quality sequence stop: 294.							
FEATURES		Location/Qualifiers		1..294		/organism="Homo sapiens"			
source		/mol_type="mrna"				/db_xref="taxon:9606"			
		/dev_stage="Adult"				/clone_lib="NT0028"			
		/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 195,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."							
BASE COUNT		86 a		57 c		74 g		77 t	
ORIGIN									
Query Match		14.8%; Score 19; DB 10; Length 294;							
Best Local Similarity		100.0%; Pred. No. 15;							
Matches		19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Qy	5	AACTACATACACACCT 23							
Db	112	AACTACATACACACCT 94							
RESULT 5		H88366							
LOCUS		H88366							
		316 bp		mrna		linear		EST 11-DEC-1995	
ORIGIN									
Query Match		14.8%; Score 19; DB 14; Length 315;							
Best Local Similarity		100.0%; Pred. No. 15;							
Matches		19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Qy	53	AAACCTAAATGCTTATAA 71							
Db	59	AAACCTAAATGCTTATAA 77							
FEATURES		Location/Qualifiers		1..315		/organism="Homo sapiens"			
source		/mol_type="mrna"				/db_xref="taxon:9606"			
		/clone="UI-H-EUI-bab-o-10-0-UI"				/tissue_type="Osteoarthritic Cartilage"			
		/dev_stage="Adult"				/lab_host="DH10B (Life Technologies)"			
		/clone_lib="NCI CGAP Ctl"				/note="Organ: Knee; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ctl is a normalized cDNA library containing the following tissue(s): Osteoarthritic Cartilage The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is TGATCAGCTT.			
		TAG LIB=UI-H-EUI							
		TAG TISSUE=Osteoarthritic cartilage							
		TAG_SEQ=TGATCAGCTT							
BASE COUNT		135 a		42 c		93 t		1 others	
ORIGIN									

```

DEFINITION Yw21c06.s1 Morton Fetal Cochlea Homo sapiens cDNA clone
IMAGE:252874 3', mRNA sequence.
ACCESSION H88366
VERSION H88366.1 GI:1069945
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 316)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favellio,A., Gish,W., Hawkins
,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Travaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 257
Source: IMAGE Consortium L1NL
This clone is available royalty-free through L1NL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 540 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 257.
Location/Qualifiers
FEATURES
source
1..316
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:252874"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="Worton Fetal Cochlea"
/clone_lib="Worton Fetal Cochlea"
/notes="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI;
Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
37% of inserts <0.5 Kb, 56% 0.5-1.0 Kb, 7% >1 Kb. Uni-ZAP
XR Vector. Library constructed by N. Robertson, C. Morton.
-5' adaptor sequence: 5' GAATTCGGCAGCG 3' -3' adaptor
sequence: 5' CTCGAGTGTGTTTTTTTTTTT 3'"
BASE COUNT 144 a 45 c 52 g 75 t
ORIGIN
Query Match 14.8%; Score 19; DB 14; Length 316;
Best Local Similarity 100.0%; Pred. NO. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 AAAACCTAAATGCTTATAA 71
|||||
Db 25 AAAACCTAAATGCTTATAA 43

RESULT 6
LOCUS AA977104 376 bp mRNA linear EST 07-JUL-1998
DEFINITION Oq24b07.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1587253 3',
mRNA sequence.
ACCESSION AA977104
VERSION AA977104.1 GI:3154550
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through L1NL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1223 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 371.

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 376)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/L1NL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 495 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 312.
Location/Qualifiers
FEATURES
source
1..376
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1587253"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC4"
/notes="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT 162 a 50 c 62 g 102 t
ORIGIN
Query Match 14.8%; Score 19; DB 9; Length 376;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 AAAACCTAAATGCTTATAA 71
|||||
Db 41 AAAACCTAAATGCTTATAA 59

RESULT 7
LOCUS AI383444 391 bp mRNA linear EST 28-MAR-1999
DEFINITION te30e09.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2087464 3', mRNA sequence.
ACCESSION AI383444
VERSION AI383444.1 GI:4196225
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through L1NL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1223 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 371.

```

FEATURES  
source

Location/Qualifiers  
1. .391  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2087464"  
/lab\_host="DH10B"  
/clone\_lib="Soares NFL T'GBC S1"  
/notes="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI CGAP GC81) were mixed, and ss circles were made in vitro. Following RAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "  
BASE COUNT 163 a 51 c 63 g 114 t

Query Match 14.8%; Score 19; DB 9; Length 391;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 53 AAAACCTAAATGCTTATAA 71  
Db 53 AAAACCTAAATGCTTATAA 71

RESULT 8  
AI024541  
LOCUS 440 bp mRNA linear EST 18-JUN-1998  
DEFINITION ov38f06.x1 Soares.testis\_NHT Homo sapiens cDNA clone IMAGE:1639619  
3', mRNA sequence.  
ACCESSION AI024541  
VERSION AI024541.1 GI:3240154  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 440)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
cDNA library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 439.  
Location/Qualifiers  
1. .440  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1639619"  
/sex="male"  
/lab\_host="DH10B"  
/clone\_lib="Soares testis NHT"  
/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5',

FEATURES  
source

Location/Qualifiers  
1. .452  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:418771"  
/db\_xref="taxon:9606"  
/clone="IMAGE:46230"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares infant brain INTB"  
/notes="Organ: whole brain; Vector: Lofmid BA; Site 1: Not I; Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',  
AACTGAGAGATTCCGCCGCCGAGAAATTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lofmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 104 a 98 c 147 g 98 t 5 others

TGTTACCAATCTCAAGTGGAGCGCGCCCAATTTTTTTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cots5, and was constructed by Bento Soares and M. Fatima Bonaldo. "  
BASE COUNT 174 a 67 c 75 g 124 t

Query Match 14.8%; Score 19; DB 9; Length 440;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 53 AAAACCTAAATGCTTATAA 71  
Db 45 AAAACCTAAATGCTTATAA 63

RESULT 9  
H09696/c  
LOCUS 452 bp mRNA linear EST 23-JUN-1995  
DEFINITION Y199e05.x1 Soares infant brain INIB Homo sapiens cDNA clone IMAGE:46230 5', similar to SP:KAD2 RAT P29410 ADENYLATE KINASE ISOENZYME 2, MITOCHONDRIAL ;, mRNA sequence.  
ACCESSION H09696  
VERSION H09696.1 GI:874518  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 452)  
AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M., Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M., Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F., Trevaskis E., Waterston R., Williamson A., Wohlmann P. and Wilson R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
Insert Size: 467  
High quality sequence stops: 335  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 467 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 335.  
Location/Qualifiers  
1. .452  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:418771"  
/db\_xref="taxon:9606"  
/clone="IMAGE:46230"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares infant brain INTB"  
/notes="Organ: whole brain; Vector: Lofmid BA; Site 1: Not I; Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',  
AACTGAGAGATTCCGCCGCCGAGAAATTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lofmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 104 a 98 c 147 g 98 t 5 others

## ORIGIN

Query Match 14.8%; Score 19; DB 14; Length 452;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 AGGGAAGCCATCCAGAG 124  
 |||||  
 Db 353 AGGGAAGCCATCCAGAG 335

## RESULT 10

AA205389  
 LOCUS zq78h02.s1 StrataGene hNT neuron (#937233) Homo sapiens cDNA clone  
 DEFINITION IMAGE:647763 3', mRNA sequence.

ACCESSION AA205389  
 VERSION AA205389.1 GI:1803380  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,  
 Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,  
 J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,  
 White, Y., Wylie, T., Waterston, R. and Wilson, R.  
 WashU-NCI human EST Project

## TITLE

## JOURNAL

## COMMENT

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LML; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1079 Std Error: 0.00

Seq primer: -40W13 fwd. from Amersham

High quality sequence stop: 321.

## FEATURES

## source

1..482  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:5218015"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:647763"  
 /dev\_stage="hNT neurons"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /clone\_lib="StrataGene hNT neuron (#937233)"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; Cloned unidirectionally. Primer: Oligo dT.  
 Differentiated, post mitotic hNT neurons. Average insert  
 size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'  
 GAATTCGCGACGAG 3' ~3' adaptor sequence: 5'  
 CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 180 a 79 c 93 g 123 t

## ORIGIN

Query Match 14.8%; Score 19; DB 9; Length 482;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 AAAACCTAAATGCTTATAA 71  
 |||||  
 Db 41 AAAACCTAAATGCTTATAA 59

## RESULT 11

AQ880140/c  
 LOCUS HS\_5037\_B1\_A10\_T7 RPCI-11 Human Male BAC Library Homo sapiens  
 DEFINITION

## ACCESSION

AQ880140

## VERSION

AQ880140.1 GI:6311607

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

**TITLE** National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index

**JOURNAL** Unpublished

**COMMENT** Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 772 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 403.  
Location/Qualifiers  
1. .497  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2292082"  
/tissue\_type="tumor, 5 pooled (see description)"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Brn52"  
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; This library represents the normalized version of NCI CGAP Brn35. Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.19 Kb. Tumor types include: meningioma, oligodendroglioma, astrocytoma (grade II), medulloblastoma, astrocytoma (grade IV). Constructed by Life Technologies."

**BASE COUNT** 197 a 76 c 86 g 138 t

**ORIGIN**

Query Match 14.8%; Score 19; DB 9; Length 497;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 53 AAAACCTAAATGCTTATAA 71  
|||||

**Db** 41 AAAACCTAAATGCTTATAA 59  
|||||

**RESULT 13**

**AW000981**

**LOCUS** wr91b05.x1 NCI\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2495025 3', mRNA sequence.

**DEFINITION** EST. 08-MAR-2000

**ACCESSION** AW000981

**VERSION** AW000981.1 GI:5847897

**KEYWORDS** EST.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** 1 (bases 1 to 504)

**AUTHORS** Bakyota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**TITLE** NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

**JOURNAL** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

**COMMENT** Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 1380 Std Error: 0.00

**FEATURES** source

1. .504  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2495025"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Kid11"  
/notes="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

**BASE COUNT** 197 a 77 c 89 g 141 t

**ORIGIN**

Query Match 14.8%; Score 19; DB 9; Length 504;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 53 AAAACCTAAATGCTTATAA 71  
|||||

**Db** 42 AAAACCTAAATGCTTATAA 60  
|||||

**RESULT 14**

**AW099685**

**LOCUS** RPCI-23-356H7.TJ RPCI-23 Mus musculus genomic clone RPCI-23-356H7, genomic survey sequence.

**DEFINITION** GSS.

**ACCESSION** AQ999685

**VERSION** AQ999685.1 GI:7074782

**KEYWORDS** GSS.

**SOURCE** Mus musculus (house mouse)

**ORGANISM** Mus musculus

**REFERENCE** 1 (bases 1 to 570)

**AUTHORS** Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akimret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.

**TITLE** Mouse BAC End Sequences from Library RPCI-23

**JOURNAL** Unpublished

**COMMENT** Other GSSs: RPCI-23-356H7.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: shao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.html) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
Plate: 356 row: H column: 7  
Seq primer: SP6  
Class: BAC ends.

**FEATURES** source

1. .570  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"



```

/clone="RPCI-23-356H7"
/sex="Female"
/lab host="DH10B"
/clone lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      213 a      101 c      84 g      172 t
ORIGIN

```

```

Query Match      14.8%; Score 19; DB 28; Length 570;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      53 AAAACCTAAATGCTTATAA 71
      |||||
Db      134 AAAACCTAAATGCTTATAA 152

```

```

RESULT 15
AZ001717
LOCUS      583 bp      DNA      linear      GSS 24-FEB-2000
DEFINITION RPCI-23-340J21.TU RPCI-23 Mus musculus genomic clone RPCI-23-340J21
, genomic survey sequence.
ACCESSION  AZ001717
VERSION     AZ001717.1 GI:7077025
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 583)
AUTHORS   Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
TITLE      Mouse BAC End Sequences from Library RPCI-23
JOURNAL
COMMENT

```

```

Other GSSs: RPCI-23-340J21.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(piet@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac/bac\_ends/mouse/bac\_end\_intro.html
Plate: 340 row: J column: 21
Seq primer: SP6
Class: BAC ends.

```

```

FEATURES             location/Qualifiers
     source           1..583
     organism="Mus musculus"
     mol_type="genomic DNA"
     strain="C57BL/6J"
     db_xref="taxon:10090"
     clone="RPCI-23-340J21"
     sex="Female"
     lab host="DH10B"
     clone lib="RPCI-23"
     note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the

```

```

EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      214 a      105 c      86 g      178 t
ORIGIN
Query Match      14.8%; Score 19; DB 28; Length 583;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      53 AAAACCTAAATGCTTATAA 71
      |||||
Db      134 AAAACCTAAATGCTTATAA 152

```

```

Search completed: August 14, 2003, 19:41:02
Job time : 1816 secs

```